AFRL-RH-WP-TR-2009-0097

Analysis of Eye Movements in Change Detection With Teams Using a Simulated Tactical Situation Display

Camilla C. Knott
National Research Council
500 Fifth St. N.W.
Washington D.C. 20001

W. Todd Nelson
Megan K. McCroskey
Brent T. Miller
Human Effectiveness Directorate
Warfighter Interface Division
Wright-Patterson AFB OH 45433

September 2008

Final Report for the period August 2006 to September 2008

Air Force Research Laboratory
711th Human Performance Wing
Human Effectiveness Directorate
Warfighter Interface Division
Collaborative Interfaces Branch
Wright-Patterson AFB OH 45433

NOTICE AND SIGNATURE PAGE

TECHNICAL REVIEW AND APPROVAL

AFRL-RH-WP-TR-2009-0097

THIS TECHNICAL REPORT HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION.

FOR THE DIRECTOR

//signed//
BENJAMIN KNOTT
Program Manager
Collaborative Interfaces Branch

//signed//
DANIEL G. GODDARD
Chief, Warfighter Interface Division
Human Effectiveness Directorate
711th Human Performance Wing
Air Force Research Laboratory

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
01-09-2008	Final	30 Aug 06 – 15 Sep 07

4. TITLE AND SUBTITLE
Analysis of Eye Movements in Change Detection with Teams Using a Simulated Tactical Situation Display

5a. CONTRACT NUMBER
In-House

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
62202F

6. AUTHOR(S)
Camilla C. Knott*
W. Todd Nelson**
Megan K. McCroskey**
Brent T. Miller**

5d. PROJECT NUMBER
7184

5e. TASK NUMBER
08

5f. WORK UNIT NUMBER
71840872

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
National Research Council*
500 Fifth St. N.W.
Washington D.C. 20001

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Materiel Command**
Air Force Research Laboratory
711th Human Performance Wing
Human Effectiveness Directorate
Warfighter Interface Division
Collaborative Interfaces Branch
Wright Patterson AFB OH 45433-7022

10. SPONSOR/MONITOR'S ACRONYM(S)
711 HPW/RHCP

11. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFRL-RH-WP-TR-2009-0097

12. DISTRIBUTION AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

88 ABW/PA Cleared 09/25/2008 88ABW-08-0300.

14. ABSTRACT
Change blindness results from the absence of attention from the source of a change. Unsurprisingly, change blindness worsens in dual task and high workload conditions. Such is the environment in which teams of operators in military command and control missions operate, yet little is known about change blindness susceptibility in these conditions. A flicker task required individuals and dyads to detect changes under high and low task difficulty. Reaction time, accuracy, eye movements, and three workload measures were examined. Scan path analyses showed that communicating dyads were more efficient at change detection than individuals and non-communicating dyads, but are just as blind to no change. Eye movement data also indicate differences in attention allocation and scan strategies among individuals, communicating, and non-communicating dyads. Implications for adaptive aids and tasks such as tactical command and control are discussed.

15. SUBJECT TERMS
Automation, Battle Management, Command and Control System

16. SECURITY CLASSIFICATION OF: Unclassified			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					Benjamin Knott
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPONE NUMBER (*Include area code*)**
U	U	U	SAR	91	

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

EXECUTIVE SUMMARY 1
INTRODUCTION 3
METHODS 7
Participants *7*
Design *8*
Stimuli & Apparatus *9*
Stimuli 9
Eye Tracker 10
Head Tracker 11
Heart Rate Monitor 11
Workload Questionnaires 12
Exit Questionnaire 12
Procedure *12*
RESULTS 14
Comparisons between groups *15*
Response Time 16
No Comm/Comm 16
Individuals/No comm 16
Individuals/Comm. 16
Accuracy 18
No Comm/Comm 18
Individuals/No comm. 18
Individuals/Comm 18
Fixations 22
No comm/Comm. 22
Individuals/No comm. 22
Individuals/Comm. 22
Cross-scanning 25
No Comm/Comm. 25
Re-scanning 27
No comm/Comm 27
Individuals/No comm. 27
Individuals/Comm. 27
Workload: NASA RTLX 29
No comm/Comm 29
Individuals/No comm 30
Individuals/Comm 30
Workload: NASA TLX subscale ratings 30
No comm/Comm, ... 30
Individuals/No comm 31
Individuals/Comm. 31
Workload: Heart Rate 34
Workload: NNI 34
No comm/Comm 34
Individuals/No comm. 34
Individuals/Comm 34
Workload: Team Workload Scale 35
Confidence Ratings 35
NNI as a Workload Measure *37*
Ceiling effect and confounding factor on NNI analysis 37
RTLX and set size 39
NNI and RTLX 39
NNI, heart rate, and RTLX 39
Regression model of change detection performance with multiple observers *39*
Response time model 39
Accuracy model 40
DISCUSSION 41
Are effects of communication on change detection performance apparent in eye movements and scan paths? *41*

Is the NNI a viable workload index? *45*
REFERENCES 47
APPENDIX A: CONSENT FORM AND BIOGRAPHICAL QUESTIONNAIRE 50
APPENDIX B: TASK INSTRUCTIONS 60
APPENDIX C: QUESTIONNAIRES 64
APPENDIX D: ANALYSIS OF ACCURACY FROM *INDIVIDUALS* AND *COMM DYADS* WITH AND WITHOUT PARTICIPANT Q 70
APPENDIX E: DATA 75

List of Figures

Figure 1. Flicker task used in this study. The top panel (A) enotes a large change, set size 48 trial and the bottom panel (B) denotes a large change, set size 6 trial. 10
Figure 2. (A)Two ASL head- mounted eye trackers were used to record eye movements. (B) A divider prevented observers from viewing each others' screens. 14
Figure 3. RT as a function of change magnitude for set size 6 and 48 (set size x change magnitude interaction, p<.05). Each graph represents results from ANOVAs with the between-subjects factor indicated above the graph. Striped bars represent set size 6 and solid bars represent set size 48. 17
Figure 4. RT as a function of set size for Individuals and Comm dyads (set size x group interaction, p<.05). Unfilled bars represent Individuals and solid bars represent Comm dyads. 17
Figure 5. Acuracy as a function of set size for No comm and Comm dyads (set size x group interaction, p<.05). Striped bars represent No comm dyads and solid bars represent Comm dyads. 19
Figure 6. Accuracy as a function of change magnitude for set size 6 and 48 (set size x change magnitude interaction, p<.05). Each graph represents results from ANOVAs with the between-subjects factor indicated above the graph. Striped bars represent set size 6 and solid bars represent set size 48. 19
Figure 7. Accuracy as a function of set size for Individuals and Comm dyads (group x set size interaction, p<.05). Unfilled bars represent Individuals and filled bars represent Comm dyads. 20
Figure 8. Accuracy as a function of change magnitude for Individuals and Comm dyads (group x change magnitude interaction, p<.05). Unfilled bars represent Individuals and solid bars represent Comm dyads. 20
Figure 9. Number of fixations as a function of change magnitude for for set size 6 and 48 (set size x change magnitude interaction, p<.05). Each graph represents results from ANOVAs with the between-subjects factor indicated above the graph. Striped bars represent set size 6 and solid bars represent set size 48. 23
Figure 10. Number of fixations as a function of change magnitude for No comm and Comm dyads for set size 6 (on the left) and 48 (set size x change magnitude x group interaction, p<.05). Striped bars represent No comm dyads and solid bars represent Comm dyads. 23
Figure 11. Number of fixations as a function of set size for Individuals and No comm dyads (set size x group interaction, p<.05). Unfilled bars represent Individuals and striped bars represent No comm dyads. 24
Figure 12. Number of fixations as a function of change magnitude for Individuals and No comm dyads for set size 6 (on the left) and set size 48 (set size x change magnitude x group interaction, p<.05). Unfilled bars represent Individuals and striped bars represent No comm dyads. 24
Figure 13. Percentage of cross-scanned trials as a function of set size for No comm and Comm dyads (set size x group interaction, p<.05). Solid bars represent Comm dyads and striped bars represent No comm dyads. 25
Figure 14: Scan paths for Individuals, No comm and Comm dyads for no change and change trials. The two-toned square is the square which changed in position. Solid lines in the Dyad trials represent the scan path of the observer that responded. 26
Figure 15. Percentage of cross-scanned trials as a function of change magnitude for No comm and Comm dyads (change magnitude x group interaction, p<.05). Striped bars represent No comm dyads and solid bars represent Comm dyads. 26
Figure 16. Percentage of re-scanned trials as a function of change magnitude for set size 6 and 48 (set size x change magnitude interaction, p<.05). Each panel represents results from ANOVAs with the between-subjects factor indicated above the graphs. Striped bars represent set size 6 and solid bars represent set size 48. 28
Figure 17: Percentage of re-scanned trials as a function of set size for No comm and Comm dyads (set size x group interaction, p<.05). Striped bars represent No comm dyads and solid bars represent Comm dyads. 28

Figure 18. Percentage of re-scanned trials as a function of change magnitude, averaged across set size (change magnitude x group interaction, p<.05). Unfilled bars represent Individuals and striped bars represent No comm dyads. 29
Figure 19. Percentage of re-scanned trial as a function of change magnitude for Individuals and Comm dyads (change magnitude x group interaction, p<.05). Unfilled bars represent Individuals and solid bars represent Comm dyads. 29
Figure 20. TLX ratings as a function of subscale for set size 6 and 48, averaged across Comm and No comm dyads (set size x subscale interaction, p<.05). 32
Figure 21. TLX ratings as a function of subscale for Comm and No comm dyads, averaged across set size (group x subscale interaction, p<.05). 32
Figure 22. TLX rating as a function of subscale for set size 6 and 48, collapsed across Individuals and No comm dyads (group x subscale interaction, p<.05). 33
Figure 23. TLX rating as a function of subscale for set size 6 and 48, collapsed across Individuals and Comm dyads (group x subscale interaction, p<.05). 33
Figure 24. NNI as a function of set size 6 and 48 for No comm and Comm dyads (group x set size interaction, p<.05). Striped bars represent No comm dyads and solid bars represent Comm dyads. 34
Figure 25. NNI as a function of set size 6 and 48 for Individuals and Comm dyads (group x set size interaction, p<.05). Unfilled bars represent No comm dyads and solid bars represent Comm dyads. 35
Figure 26. Average dyad rating as a function of team workload demand subscale for set size 6 and 48 (set size x subscale interaction, p<.05). Striped bars represent set size 6 and solid bars represent set size 48. 36
Figure 27. (A) Scan paths for Individuals and No comm dyads for trials with No change, Small and Large change. Scan paths of Individuals are shown against a gray background. (B) A greater portion of No change displays were scanned compared to Small and Large change displays. (C) Scan paths for Comm dyads for trials with No change, Small, and Large change. 45

List of Tables

Table 1. List of participants and counterbalance order. Only data from participants who participated in both the individual and dyad conditions were included in the analysis (Italicized). Dyads consisted of individuals with an equal amount of experience with the task and the same counterbalance order, e.g., Dyad 6 consisted of two individuals with 2 sessions of prior experience with the task, once as an individual and once as part of a dyad. 8
Table 2. Means and standard errors for RT, accuracy, fixations, % of re-scanned and cross-scanned trials for Individuals, Comm, and No comm dyads for set size 6 and set size 48. 211
Table 3. Means and SEs for RTLX, heart rate, and NNI for Individuals, No comm and Comm dyads and Team workload scale (TWS) for Comm dyads. 300
Table 4. Regression and ANOVA tables for the model of accuracy in change detection with multiple observers. 400

Executive Summary

This post doctoral research project was aimed at following up on a dissertation project by Dr. Allison Tollner in which she investigated change blindness in individuals and teams. The primary goal was to examine the role of attention and eye movements on the ability of individual and multiple observers to detect changes in a complex visual scene. Related to this goal, the project involved the set up of an eye tracking laboratory which allowed simultaneous recording of eye movements from dyads. A secondary goal was to determine the viability of the Nearest Neighbor Index, an analysis of point dispersion in space, as a physiological measure of cognitive workload in a change detection task.

Change blindness results from the absence of attention from the source of a change. Unsurprisingly, change blindness worsens in dual task and high workload conditions. Such is the environment in which teams of operators in military command and control missions operate, yet little is known about change blindness susceptibility under these conditions. A flicker task required individuals and dyads to detect changes under easy and difficult task conditions. Response time, accuracy of responses, eye movements, heart rate, and subjective workload measures were recorded and analyzed.

Tollner's findings indicated that change detection performance only benefits from having more eyes or resources when observers are allowed to communicate and thereby coordinate their actions and strategies. Results of the present study replicate these findings by showing that communicating dyads were more efficient at change detection than individuals and non-communicating dyads. Our analyses of eye movements provide support for the importance of communication in mitigating change blindness susceptibility in teams of observers.

Overall, scan paths and eye movement data show differences in attention allocation and scan strategies among individuals, communicating, and non-communicating dyads. Our data show that communicating dyads functioned as a team, with several major components of teamwork evident in their scan paths. This likely contributed to more efficient change detection strategies and therefore, better performance overall. In contrast, non-communicating dyads appeared to depend on slow and inefficient search strategies in comparison to individual observers and communicating teams. In addition, communicating dyads are just as blind to no change as individual observers and non-communicating dyads. All groups, regardless of ability

to communicate or number of eyes on task, were slowed to determine the absence of change in the visual scene and yielded less efficient eye movement patterns on trials in which no change occurred compared to those in which a change did occur. Such behaviors could delay decision-making in environments such as tactical command and control and lead to catastrophic outcomes. These findings have implications for adaptive aids and tasks such as tactical command and control in which teams of operators depend on a common and current understanding of the battlespace for decision-making.

Results of the Nearest Neighbor Index (NNI) analysis suggest that the NNI may be a viable measure of mental workload for change detection paradigms, but within limits. First, given that NNI is based on an average distance between points, a major limitation of the measure is the necessity for enough points for accurate and reliable NNIs. This limitation may preclude the use of NNI to measure workload in time limited tasks. In the present study, fixations from multiple trials and change magnitude conditions were combined to overcome this problem and obtain a reliable NNI. However, the use of NNI may be constrained in environments such as tactical C2 in which workload may vary within relatively short periods of time yielding only few fixations for NNI calculation. Second, our experimental design required the randomization of icon positions which confounded the calculation of the NNI. In the easy condition, fewer stimuli were scattered in a larger space which resulted in greater variability in distances between points in the easy compared to the difficult condition. Consequently, the average distance between points was closer to random in the easy compared to the difficult condition, leading to higher NNIs in the easy compared to the difficult condition. This is counter to previous findings of higher NNI values with more difficult tasks and our results of poorer performance in the difficult compared to the easy condition.

Two regression models were developed to determine the best predictors of change detection performance of dyads.

Introduction

Change Blindness (CB) is the lack of awareness of considerable changes in a visual scene when attention is diverted from the source of the change by a simultaneously occurring event, e.g., screen flashes (Rensink, O'Regan & Clark, 1997). CB results from the absence of visual attention from the source of the change (Simons & Levin, 1997). Unsurprisingly, CB has been shown to increase when two or more tasks are performed simultaneously (McCarley, Vais, Pringle, Kramer, Irwin, & Strayer, 2004), and when a secondary task imposes high demands on attention (Muthard & Wickens, 2003). One domain that may be particularly relevant to the CB phenomenon is tactical command and control (C2). This task requires teams to monitor and maintain shared awareness of the tactical battlespace and immediately detect and communicate critical changes, e.g. sudden disappearance of aircraft from the display. Failure to do so could prove catastrophic.

Since CB occurs due to an observer's inattention to a change source, it could be argued that multiple observers would be less susceptible to CB than individual observers. Essentially, more eyes on a scene should equal more resources that may be allocated to the task. Consistent with this idea, Tollner, Riley, Nelson, Shockley, and Cummins-Sebree (2006) found that groups of observers were less susceptible to CB than individuals, but only if they communicated verbally. Tollner and her colleagues used a flicker task in which they manipulated task difficulty and the ability of observers to verbally communicate. Results of that study suggest that having more eyes or resources for change detection are only beneficial if there is a mechanism such as verbal communication for coordinating the allocation and application of those resources, thereby allowing multiple observers to function as a team (Dickinson & McIntyre, 1997). Such a conclusion may be predicated with direct evidence from visual scanning behavior. However, since Tollner's methodology did not include eye movements, the effects of communication on attention allocation and scanning strategies in teams are unknown.

The present study was designed to extend the work of Tollner et al. (2006) by investigating attention allocation and the role of scanning strategies in CB susceptibility in individuals and multiple observers. We examined attention allocation through fixations and strategies through scanning patterns. Additionally, we sought to examine the effectiveness of the Nearest Neighbor Index (NNI) in serving as a physiological measure of mental workload.

Attention allocation was examined by analysis of the number of fixations required before a response was made. This measure was based on two assumptions: saccadic eye movements executed to foveate stimuli (1) precede fixations, and (2) are obligatorily linked with shifts of attention (Shephard, Findlay, & Hockey, 1987). However, since attention may be shifted in the absence of saccades, i.e., more than one attentional shift may occur during a fixation (Posner, 1980; Shephard, et al., 1987), the measure is merely an estimate of shifts or re-allocation of attention during a task.

Scanning efficiency was measured by analysis of the number of trials with re-scanning or cross-scanning. A re-scanned trial was a trial in which (1) fixations landed on a stimulus more than once, or (2) fell within 1 degree (°) of visual angle of each other, and (3) two or more scan paths fell within 1° of visual angle of each other. The 1° of visual angle re-scan threshold was selected based on evidence that the attentional focus, the functional region of the visual field wherein information is selected for immediate perceptual processing, may be scaled to the size of relevant stimuli (Castiello & Umilta, 1990; Eriksen & Yeh, 1985). In the present study, stimuli were approximately 1° of visual angle in size and their movements were either .5° or 1.5° of visual angle, or 1° on average. A cross-scanned trial was a trial in which scan paths of partners in a dyad overlapped.

Re-scanning and cross-scanning are based on the re-checking hypothesis proposed by Treisman and Gelade (1980) to account for findings of longer search times in visual search when the specified target was absent compared to when it was present in the array. They argued that, on target absent trials, people not only search the stimulus set exhaustively, they also re-check stimuli for confirmation of their identities. Consequently, this inefficient search strategy slowed responses on target absent trials. Similarly, we assumed that a greater proportion of re-scanned and cross-scanned trials represented an inefficient search for change, or change detection strategy.

A secondary goal of this study was to assess the efficacy of the NNI as a physiological measure of cognitive workload. Workload is defined as the mental effort or cognitive processing resources required for performing a task (Norman & Bobrow, 1975), with difficult tasks typically leading to high workload by placing higher demands on a finite amount of cognitive processing resources. Eye movements have been used to measure workload with varying results (reviewed in DiNocera, Terenzi, & Camilli, 2006). In the

present study, we used NNI to relate fixation distribution to workload changes. NNI is derived by comparing the mean distance between actual points (e.g., fixations) within a specified area to the mean distance given a theoretical random distribution of the same number of points in an equal area (Clark & Evans, 1954). Given that NNI is the ratio of the average distance between points in an actual and a theoretical random distribution, the more random the distribution, the closer the NNI value will be to one.

Recently, DiNocera, Camilli & Terenzi (2007) used the NNI to measure pilot workload, and confirmed previous results from DiNocera, Terenzi, & Camilli (2006) that NNI, thus, fixation randomness, increased with workload. Consistency across two different tasks – one used static and one used dynamic displays - suggests that NNI may provide a domain-independent workload measure. NNI may be calculated easily, frequently during a task, and with data from affordable, non-invasive technology, making it accessible to many researchers. Therefore, the viability of NNI as a measure of workload should be further investigated.

CB has been shown to increase with workload (Muthard & Wickens, 2003), which varies significantly during C2 tasks such as Air Battle Management (Knott, Bolia, Nelson, & Galster, 2006). Where catastrophic outcomes may follow workload-related performance degradations, timely assessment of workload may allow for (a) determination of changes in task distribution within teams, or (b) triggering adaptive automation aids to curtail performance degradations (Parasuraman, 2003). However, most workload measures require data to be gathered over time to determine physiological and behavioral responses to changes in task requirements, or a task is interrupted to administer subjective measures such as the National Aeronautics and Space Administration Task Load Index (NASA TLX: Hart & Staveland, 1988). Additionally, both physiological and behavioral measures have provided conflicting results across tasks. Therefore, multiple workload measures are recommended for confirmation of results within a study (Wickens, Gordon, & Liu, 1998).

In the present study, we examined the efficacy of NNI as an immediate and reliable workload measure. To this end, we analyzed heart rate, and subjective TLX ratings in conjunction with NNI for the purpose of comparing the results for consistency. Heart rate is the most widely used cardiac measure of workload and has been shown to reliably increase with workload (e.g., Wilson & Eggemeier, 1991; Wilson, 2002). Subjective measures such

as the NASA TLX are commonly used in applied science settings (e.g., air traffic management), and has also been shown to reliably increase with workload.

In addition, we used the Team Workload Scale (TWS: Hildebrand, Pharmer, & Weaver, 2003) to obtain a subjective measure of the workload experienced by a team of observers and to understand the role of communication on subjective workload. Given that individual workload in is defined by task demands on a limited amount of processing resources, team workload is defined as demands of a task on a limited amount of resources available from a team of individuals to perform both individual and team-related tasks such as communicating and coordinating actions (Bowers, Braun, & Morgan, 1997). However, the TWS has not been experimentally validated and its reliability is unknown. This questionnaire was used in this study (a) to assess the contributions of verbal communication in a change detection task with multiple observers and (b) to maintain consistency between methods used in the present study and that of Tollner et al (2006).

We investigated how attention allocation and visual scanning strategies in change detection vary when: (1) individuals and teams perform the task; (2) communication is allowed or prohibited; and (3) task difficulty is high or low. Additionally, we compared results of the NNI measure to other measures of workload to assess its effectiveness in capturing changes in workload on a change detection task.

Methods

Participants

Twenty participants were paid $15/hr for participation, and tested as individual observers and as part of a dyad. At the beginning of each session, each participant was required to complete an informed consent form and a biographical questionnaire (see Appendix A). Only those who participated in both the individual and dyad conditions were included in the final analysis (3 males & 6 females with mean age of 23.4 years). Participants were excluded based on two criteria: (1) inability to return for subsequent testing; (2) eye tracker calibration difficulties due to sleepiness or glossiness of eyes. All but one participant's performance data fell within 2.5 SD from the mean - participant Q's average accuracy for the large change set size 48 condition fell at 2.66 SD from the mean. Results from ANOVAs with and without participant Q's accuracy data were compared and all but one interaction were the same. The ANOVA comparing individual observers and communicating dyads without participant Q's data yielded a set size x group interaction which was only marginally significant (p = .07) with participant Q's data included in the analysis (see Appendix D for analysis of accuracy from individuals and communicating dyads with and without participant Q). Given the minimal difference in the results with and without Q's data and the benefits of including Q's data from both the individual observer and dyad conditions, participant Q's data were included in all analyses.

Participants were tested as individuals and as part of a dyad in different sessions, in counterbalanced order. As shown in Table 1, all but 1 of the participants included in the analyses were tested twice in the dyad condition, each time with a different partner having the same counterbalance assignment and experience with the task. Participant Q was tested only once as part of a dyad. A total of 11 dyads were tested in both a Communication and a No communication condition. Two dyads were excluded due to inability of a member of the dyad to return for testing in the Individual observer condition. One dyad was excluded due to eye movement recording error in which data were lost. Also shown in Table 1, of the remaining participants whose data were included in the analyses, half were tested first as individuals and half were tested last as individuals. A within-subjects design was used to control for potential baseline performance differences when individuals and dyads perform

the same task. However, as described in the Results section, a between-subjects factor was used in comparing data from individuals and dyads.

Table 1. List of participants and counterbalance order. Only data from participants who participated in both the individual and dyad conditions were included in the analysis (Italicized). Dyads consisted of individuals with an equal amount of experience with the task and the same counterbalance order, e.g., Dyad 6 consisted of two individuals with 2 sessions of prior experience with the task, once as an individual and once as part of a dyad.

PARTICIPANT	SESSION 1	SESSION 2	SESSION 3
A	Individual 1	No dyad	
B	***Individual 2***	***Dyad 3***	***Dyad 6***
C	***Individual 3***	***Dyad 4***	***Dyad 6***
D	Individual 4	No dyad	
E	Individual 5	No dyad	
F	Individual 6	No dyad	
G	***Individual 7***	***Dyad 3***	***Dyad 8***
H	Individual 8	No dyad	
I	Individual 9	No dyad	
J	***Individual 10***	***Dyad 4***	***Dyad 8***
K	Callibration problems / No dyad		
L	Callibration problems / No dyad		
M	Dyad 2 / Lost data	***Dyad 11***	***Individual 11***
N	***Dyad 5***	***Dyad 11***	***Individual 12***
O	Dyad 2 / Lost data	***Dyad 9***	***Individual 13***
P	Dyad 7 / No Individual from partner	***Dyad 9***	***Individual 14***
Q	***Dyad 5***	***Individual 15***	
R	Dyad 1 / No Individual from partner	Individual 16	
S	Dyad 1 / No Individual from partner	No individual	
T	Dyad 7	No individual	

Design

The design of this study is based directly on Tollner et al. (2006) design and the same task was used in both studies. However, we used only a subset of the levels of each independent variable used in Tollner et al. (2006) study. Two between-subjects factors with two levels each (1) Group: *Individual* or *Dyad,* and (2) Communication: *Communication* or *No communication* yielded 3 conditions: *Individual*, *Communication* (*Comm*) *dyads, and No communication (No comm) dyads.* These factors were not completely crossed, as the Communication factor applied only to *Dyads*. Within-subjects factors were (1) Set size with 2 levels: *6* or *48* icons, and (2) Change magnitude with 3 levels: *no*, *small* or *large change*.

Stimuli & Apparatus

Stimuli

Stimuli were created with custom scripts written in JAVA version 1.4.2. Each display consisted of an even number of red, blue, and green squares which were approximately 1° of visual angle in size (i.e., squares were 1.4 cm^2 each viewed from approximately 72 cm) and randomly distributed across the entire screen. Display set size was *6* or *48* squares and blocked. As depicted in Figure 1, a trial consisted of four images presented in the following order and duration: display A (750 ms), mask (250 ms), either display A, or A' with one square changed in position (750 ms), mask (250 ms). The sequence repeated until the participant responded. A block consisted of 36 trials: 12 *no change* and 24 *change* trials with one square changed in position. A 0.5° visual angle movement in any direction was a *small change*, and a 1.5° visual angle movement was a *large change*. Change magnitude was mixed within each block and randomly presented. A Pentium 4 desktop computer with 2 GB of RAM was used to present the displays on an 18” NEC Multisync LCD 1800 flat screen monitor. In the *Dyad* condition, two independent computers were used to present the same displays on two separate monitors.

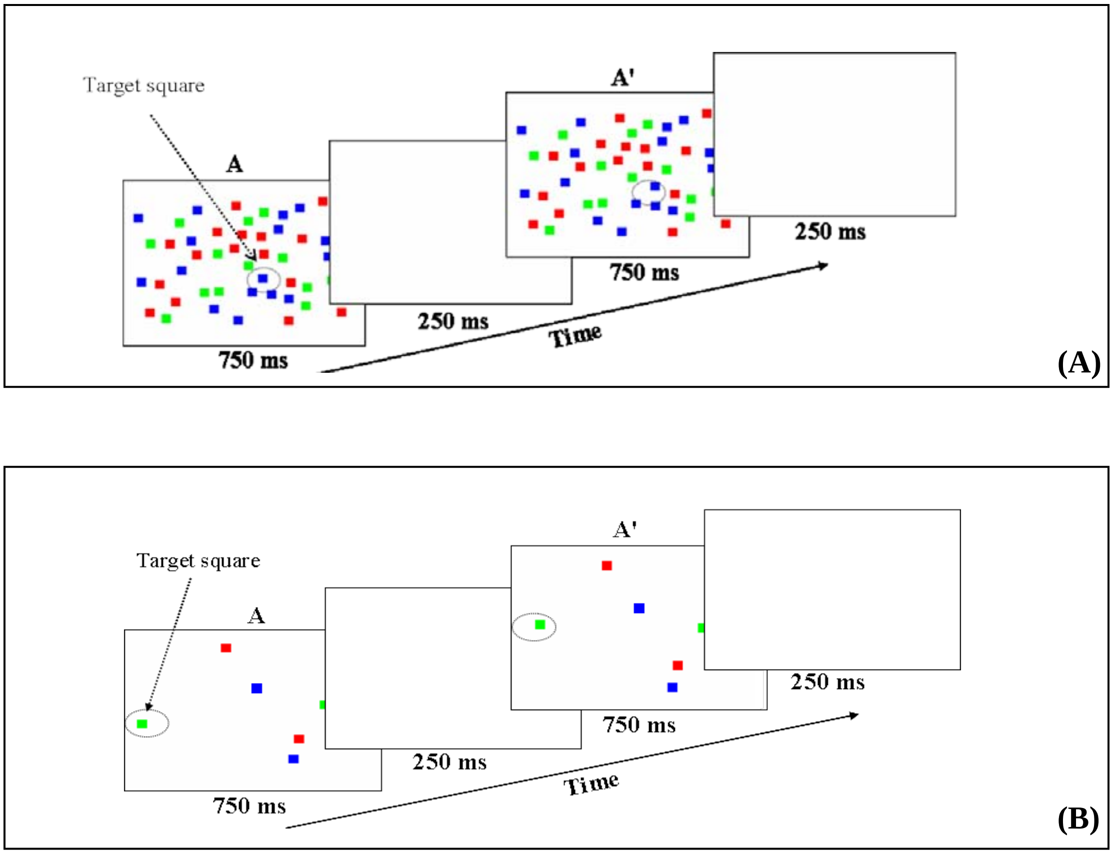


Figure 1. Flicker task used in this study. The top panel (A) enotes a large change, set size 48 trial and the bottom panel (B) denotes a large change, set size 6 trial.

Eye Tracker

Two ASL model 501 head-mounted eye trackers with head tracking were used to record eye movements (Figure 2). This system measures infrared corneal reflectance and pupil position to determine point of gaze. Fixations were points of gaze with maximum change of 1° visual angle for a minimum of 100 ms. Eye movements were recorded at 120 Hz for all but four participants in the *Individual* condition whose eye movements were recorded at 240 Hz. A mixed model ANOVA, with sampling rate as the between-subjects factor, was used to compare performance and number of fixations from these two groups of participants. No significant group effects or interactions emerged (all p>.05), therefore data from the two groups were combined for further analyses.

Eye tracker error resulted in artificial fixations and fixations which fell outside of the display boundaries. These fixations were filtered out and excluded from the analyses. Artificial fixations were fixations that were preceded by saccadic eye movements with speeds greater than 700° visual angle/sec. The eye tracker system is programmed to calculate point of gaze using head position, pupil location, and corneal reflection (CR). It expects the CR created by the infrared light source on the head-mounted camera to be the brightest spot on the eye. However, in some cases, reflections from other light sources in the testing room (e.g., monitors) or glossy parts of the eye created spots that were at least as bright as the CR. This caused the system to oscillate between tracking the actual CR and the brighter spot, thereby shifting point of gaze. Consequently, artificial fixations were created whenever point of gaze shifted more than 1° of visual angle for a period exceeding 100 ms.

Head Tracker

A Flock of Birds magnetic tracker was used to track head movement. Each eye tracker was connected to its own control unit, with the two control units configured in master/slave mode. However, as shown in figure 2, both eye trackers shared one magnetic transmitter which was positioned between the two stations, approximately 3 feet from each participant. The Flock of Birds magnetic field ranges approximately 4 feet and samples for changes in head motion at approximately 144 Hz (http://www.ascension-tech.com/products/flockofbirds.php). An eye/head integration mode on the ASL eye tracker allowed the eye tracker to subtract head movements from point of gaze calculations.

Heart Rate Monitor

A UFI EZ-IBI heart rate monitor was used to record inter-beat intervals (IBI). This system required three surface skin electrodes to be attached to the participant for detection of the electrocardiogram (ECG) signal. Two electrodes were attached to the chest: one on the sternum and one on the left side of the body below the 5^{th} rib, approximately 6 inches from the armpit. A third electrode was attached to the right wrist and served as the ground. The skin below the electrodes was prepared by applying NuPrep ECG and EKG skin abrasion gel to the site, gently wiping the skin with gauze, then cleansing the site with alcohol prior to application of the electrodes. The UFI EZ-IBI monitor recorded data at the rate of 1000 Hz.

IBIs were time periods between R wave peaks and were filtered off-line using Microsoft Access version 11 from Microsoft Office Suite 2003.

Workload Questionnaires

Participants in both the *Individual* and *Dyad* conditions completed an electronic version of the NASA TLX (Hart & Staveland, 1988) which is comprised of 6 subscales: Mental demand, Physical demand, Temporal demand, Effort, Performance, and Frustration (see Appendix C1). Each subscale required a rating between 0 and 100. Ratings from the 6 subscales were averaged to obtain a Raw TLX score for data analysis (RTLX: Byers, Bitner, & Hill, 1989). *Dyad* RTLX scores were obtained by averaging subjective workload ratings from both members (Bowers, Braun, & Morgan, 1997). Raw data, with ratings from the individual subscales, were also analyzed to assess differences in subscale ratings between groups.

In addition to the NASA TLX, *Dyads* also completed the Team Workload Scale (Appendix C2: Hildebrand, Pharmer, & Weaver, 2003). This questionnaire consisted of five subscales: Communication demands, Monitoring demands, Control demands, Coordination demands, Leadership demands, with 3 questions under each subscale. Each question required a rating between 0 and 20 with 20 indicating that the rater experienced high workload under a specified condition. Dyad subscale scores were obtained by averaging ratings from members of each dyad.

Exit Questionnaire

All participants were required to complete an exit questionnaire at the end of each testing session (Appendix C3).

Procedure

Participants were seated approximately 72 cm from the screen. In the *Dyad* condition, members viewed the same displays on two different monitors. As shown in Figure 2 a cardboard divider was used to prevent members of a dyad from viewing each other's monitors. Participants were asked to read the instructions for the task at the beginning of

each testing session (see Appendix B1 for instructions for *Individuals* and Appendix 2b for *Dyads*)

A flicker task required participants to indicate by key press if one square in an array changed in position. To discourage guessing, participants were required to indicate by mouse click which square changed in position whenever they responded that a change occurred. All participants used his or her left hand on the keyboard to indicate whether or not a change occurred ("A" for YES and "D" for NO) and his or her right hand on the mouse to indicate which square changed in position. In the *Dyad* condition, participants viewed the same displays on separate monitors. To prevent strategy carry-over, *Dyads* were tested first in the *No comm* condition in which communication was prohibited before, during, and between trial blocks. *Dyads* were then tested in the *Comm* condition in which they were encouraged to communicate and develop strategies for change detection. Participants completed a set of practice trials prior to each condition. Trials were self-paced so that the subsequent trial began only when the participant pressed the space bar to continue.

At the end of each set size block, participants were required to complete a NASA TLX questionnaire to allow us to determine perceived workload and validate the task difficulty manipulation of set size. In addition, to obtain a measure subjective perception of performance as a team, participants in the *Dyad* condition were also required to complete the Team Work Load Scale at the end of each set size block. All participants completed an exit questionnaire on strategies used and confidence in responses given during the task at the end of each testing session (Appendix C). Participants provided confidence ratings on a scale from 1 to 10 with 10 indicating the highest level of confidence in their own performance of the task.

(A) (B)

Figure 2. (A)Two ASL head- mounted eye trackers were used to record eye movements. (B) A divider prevented observers from viewing each others' screens.

Results

Three sets of analyses were performed on the data. The first set below is comprised of comparisons of data between groups, i.e., *Individuals, No comm dyads,* and *Comm dyads.* The second set focused on determining the efficacy of NNI as a measure of workload. The third set is a regression model predicting change detection performance in dyads.

Change detection performance was measured using response time (RT in ms) and accuracy (percentage correct), and attention allocation and scanning strategy were examined by analysis of fixations and scan paths. For Dyads, fixations from both members were summed for analysis to reflect the total amount of effort required from the dyad. However, each member's percentage of re-scanned trials was analyzed separately since re-scanning varied between dyad members from trial to trial, e.g., one member re-scanned on trial 9, but not trial 10 and the other member re-scanned on both trial 9 and trial 10.

Workload was assessed, physiologically, by analyses of heart rate and NNI, and subjectively, by analyses of RTLX, TLX subscale ratings, and Team Workload Scale average ratings by set size block. Average heart rate and NNI were obtained by combining data across change magnitude condition, within each set size block. For dyads, heart rate was obtained by averaging both dyad members' heart rate for each set size. Dyad NNIs were also obtained by averaging both dyad members' NNI for each set size. Confidence ratings on responses given during the task were averaged between dyad members and compared across groups in a one-way ANOVA.

Comparisons between groups

Performance data along with number of fixations, and percent of re-scanned and cross-scanned trials were submitted to mixed model ANOVAs. To determine the effect of having more eyes and communication on change detection, two separate ANOVAs were performed to compare *Individuals* to *No comm dyads* and *Individuals* to *Comm dyads*. Data were submitted to a 2 (group: *Individual/Dyad*) x 2 (set size: *6/48* icons) x 3 (change magnitude: *small/large/no change*) mixed model repeated measures ANOVA with group as the between-subjects factor. Although all participants were tested in both *Dyad* and *Individual* conditions, a between-subjects factor was used since responses were difficult to assign to a specific member in a dyad. This was particularly problematic in the *Comm* condition in which members were encouraged to collaborate to determine a response. To further understand the effect of communication on change detection, data from *No comm* and *Comm* dyads were submitted to a 2 (group: *No comm*/*Comm dyad*) x 2 (set size) x 3 (change magnitude) repeated measures ANOVA.

To determine differences in workload across groups, workload measures, i.e., RTLX, NNI and heart rate were each submitted to a 2 (set size) x 2 (group) repeated measures ANOVA for the *No comm*/*Comm dyad* analysis and a 2 (set size) x 2 (group) mixed model ANOVA with group as between-subjects factor for the *Individual/No comm dyads* and *Individual/Comm dyads* analyses. For the analysis of heart rate, no significant main effects or interactions with group emerged. Therefore, heart rate data from all groups were combined and submitted to a t-test. Additionally, to determine differences in workload between groups across TLX subscales, another set of ANOVAs were performed on the TLX ratings. NASA TLX ratings were submitted to a 2 (set size) x 6 (subscale) repeated measures ANOVA for the *No comm*/*Comm dyad* comparison and a 2 (group: *Individual/Dyad*) x 2 (set size) x 6 (subscale) mixed model ANOVA, with group as the between subjects factor, for the *Individual/No comm dyad* and *Individual/Comm dyad* comparisons. Team Workload Scale average ratings from *Comm dyads* were submitted to a 2 (set size) x 5 (subscale) repeated measures ANOVA. Confidence ratings were submitted to three independent unpaired t-tests: *No comm/Comm, Inidiv/No comm, Indiv/Comm***.** A Huynh-Feldt (1970) correction factor was used to adjust for violations of sphericity on all repeated measures ANOVAs and a Bonferroni correction was employed for all post-hoc comparisons (unless specified, alpha

level = .017 since most tests conducted required three pairwise comparisons). Means and standard errors for RTs, accuracy, fixations, and percent of re-scanned and cross-scanned trials are reported in Table 2.

Practice effects were assessed by comparison of RT and percent correct of participants tested as individuals first and last in a mixed model ANOVA, with testing order as the between-subjects variable. No significant group effects or interactions emerged (all $p>.05$).

Response Time

No Comm/Comm. Overall, RTs were slower with set size *48* than *6*, $F(1,6) = 83.75$, $p<.05$, and varied with change magnitude, $F(2,12) = 38.25$, $p<.05$. All post-hoc comparisons were significant ($p<.017$). RTs were slowest to *no change* and fastest to *large change* trials. A change magnitude x set size interaction emerged from the analysis of RT, $F(2,12) = 25.94$, $p<.05$. As shown in Figure 3a, the change magnitude effect was driven mostly by RT variations with change magnitude when set size was *48*.

Individuals/No comm. Overall, RTs were slower with set size *48* than *6*, $F(1,14) = 73.59$, $p<.05$, and varied across change magnitude $F(2,28) = 60.35$, $p<.05$. All post-hoc comparisons were significant ($p<.017$), with RTs slowest to *no change* and fastest to *large change* trials. A change magnitude x set size interaction emerged, $F(2,28) = 37.58$, $p<.05$. Figure 3b shows that the change magnitude effect was due mostly to RT variations with change magnitude when set size was *48*.

Individuals/Comm. Overall, RTs were slower with set size *48* than *6*, $F(1,14) = 66.92$, $p<.05$, and *Comm dyads* were faster than *Individuals* to detect changes, $F(1,14) = 4.98$, $p<.05$. A set size x group interaction emerged $F(1,14) = 7.56$, $p<.05$. As shown in Figure 4, the effect of group on RT was due mostly to RT differences between *Individuals* and *Comm dyads* when set size was *48,* such that *Individuals* were particularly slower than *Comm dyads* when set size *was 48*. Averaging across set size and group, RT varied across change magnitude $F(2,28) = 68.99$, $p<.05$. All post-hoc comparisons were statistically significant ($p<.017$), with RTs slowest to *no change* and fastest to *large change* trials. A change magnitude x set size interaction emerged, $F(2,28) = 45.56$, $p<.05$. As shown in Figure 3c, the change magnitude effect on RT was driven mostly by RT variations with change magnitude when set size was *48.*

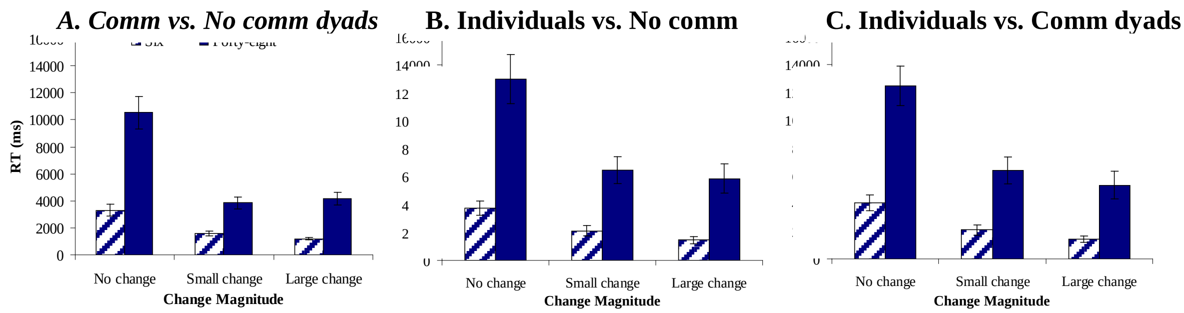


Figure 3. RT as a function of change magnitude for set size 6 and 48 (set size x change magnitude interaction, p<.05). Each graph represents results from ANOVAs with the between-subjects factor indicated above the graph. Striped bars represent set size 6 and solid bars represent set size 48.

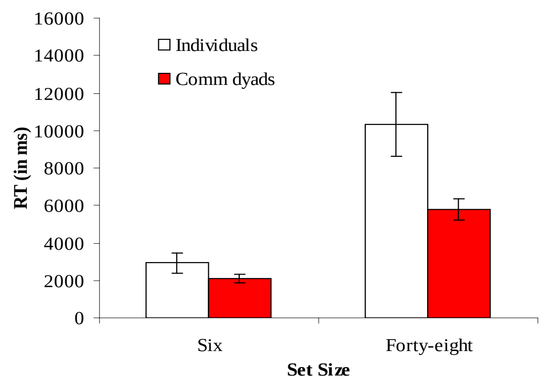


Figure 4. RT as a function of set size for Individuals and Comm dyads (set size x group interaction, p<.05). Unfilled bars represent Individuals and solid bars represent Comm dyads.

Accuracy

No Comm/Comm. Overall, accuracy was lower with set size *48* than *6*, $F(1,6) = 8.22$, $p<.05$. However, a group x set size interaction emerged from the analysis, $F(1,6) = 21.30$, $p<.05$. Figure 5 shows that *No comm dyads* were less accurate than *Comm dyads* particularly when set size was *48*. Also, overall, accuracy varied with change magnitude, $F(2,12) = 6.21$, $p<.05$. Post-hoc comparisons showed lower accuracy on *small* compared to *large* and *no change* trials, $p<.017$. A change magnitude x set size interaction emerged, $F(2,12) = 6.65$, $p<.05$. Figure 6a shows that the effect of change magnitude on accuracy was driven mostly by accuracy variations with change magnitude when set size was *48*.

Individuals/No comm. Overall, accuracy was lower with set size *48* than *6*, $F(1,6) = 8.22$, $p<.05$, and varied with change magnitude, $F(2,12) = 19.13$, $p<.05$. Post-hoc comparisons showed lower accuracy on *small* compared to *large* and *no change* trials, $p<.017$. A set size x change magnitude interaction emerged, $F(2,12) = 8.36$, $p<.05$. Figure 6b shows that the change magnitude effect on accuracy was driven by accuracy variations with change magnitude when set size was *48*.

Individuals/Comm. Overall, *Individuals* were less accurate than *Comm dyads*, $F(1,14) = 5.62$, $p<.05$. On average, both groups were less accurate with set size *48* than *6*, $F(1,14) = 9.77$, $p<.05$, and yielded accuracy variations with change magnitude, $F(2,12) = 17.45$, $p<.05$. Post-hoc comparisons of the change magnitude effect showed lower accuracy on *small* compared to *large* and *no change* trials, $p<.017$. Figure 6c depicts a set size change x magnitude interaction, $F(2,12) = 5.54$, $p<.05$: the change magnitude effect on accuracy was driven by accuracy variations with change magnitude when set size was *48*. This analysis also revealed a group x set size interaction $F(1,14) = 5.62$, $p<.05$, and a group x change magnitude interaction, $F(2,28) = 3.52$, $p<.05$. Figure 7 shows that the effect of group on accuracy was driven by accuracy differences when set size was *48* than when set size was *6*, and as shown in Figure 8, by greater accuracy variations between groups on *Small change* trials.

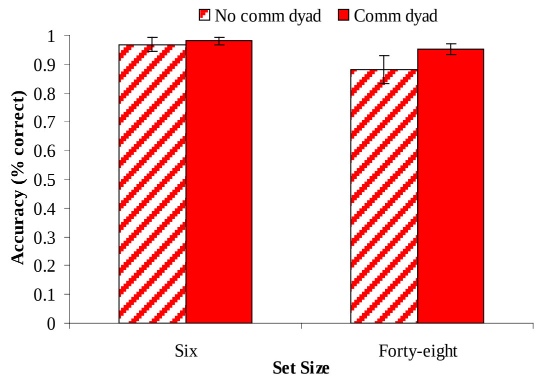


Figure 5. Acuracy as a function of set size for No comm and Comm dyads (set size x group interaction, p<.05). Striped bars represent No comm dyads and solid bars represent Comm dyads.

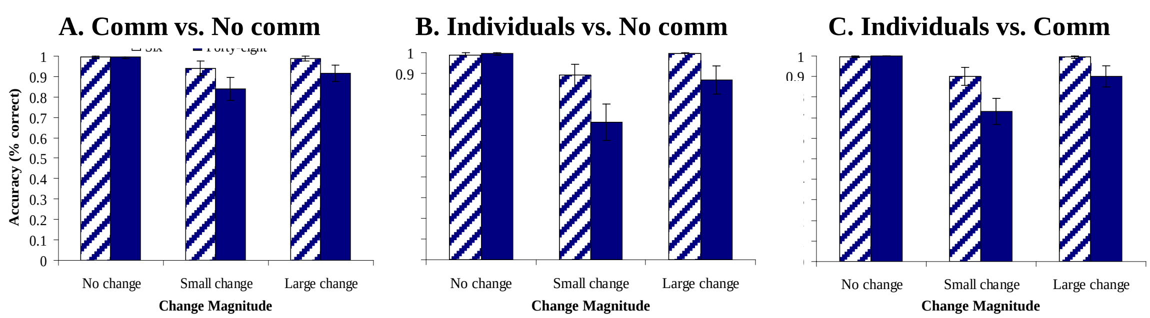


Figure 6. Accuracy as a function of change magnitude for set size 6 and 48 (set size x change magnitude interaction, p<.05). Each graph represents results from ANOVAs with the between-subjects factor indicated above the graph. Striped bars represent set size 6 and solid bars represent set size 48.

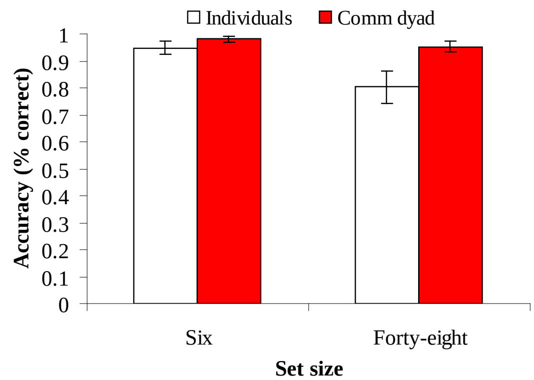


Figure 7. Accuracy as a function of set size for Individuals and Comm dyads (group x set size interaction, p<.05). Unfilled bars represent Individuals and filled bars represent Comm dyads.

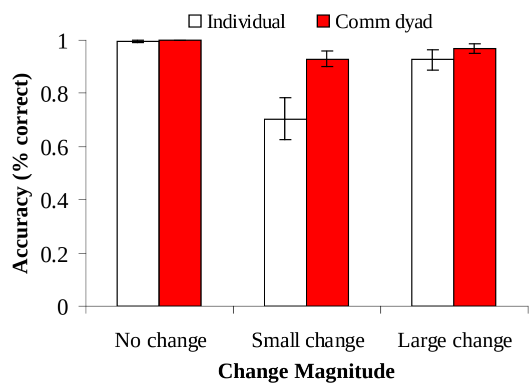


Figure 8. Accuracy as a function of change magnitude for Individuals and Comm dyads (group x change magnitude interaction, p<.05). Unfilled bars represent Individuals and solid bars represent Comm dyads.

Table 2. Means and standard errors for RT, accuracy, fixations, % of re-scanned and cross-scanned trials for Individuals, Comm, and No comm dyads for set size 6 and set size 48.

(A)

SET SIZE 6	Large change	*SE*	Small change	*SE*	No change	*SE*
RESPONSE TIME(in ms)						
Individuals	1701.72	*383.63*	2639.83	*533.37*	4457.00	*650.90*
No Comm teams	1202.21	*94.83*	1612.93	*165.90*	3015.71	*364.56*
Comm teams	1173.50	*58.38*	1535.86	*202.23*	3590.57	*490.39*
ACCURACY(% correct)						
Individuals	100%	*0.00*	85.19%	*0.06*	99.07%	*0.01*
No Comm teams	98.81%	*0.01*	92.86%	*0.05*	98.81%	*0.01*
Comm teams	98.81%	*0.01*	95.24%	*0.02*	99.99%	*0.00*
NUMBER OF FIXATIONS						
Individuals	6.92	*0.95*	7.47	*1.92*	13.38	*2.55*
No Comm teams	8.40	*0.60*	9.77	*1.25*	15.40	*2.36*
Comm teams	9.14	*0.32*	10.37	*0.69*	19.26	*1.46*
% RE-SCANNED TRIALS						
Individuals	84.26%	*0.05*	84.72%	*0.05*	89.73%	*0.05*
No Comm teams	63.15%	*0.04*	73.99%	*0.05*	87.28%	*0.05*
Comm teams	71.75%	*0.05*	77.27%	*0.05*	95.24%	*0.02*
% CROSS-SCANNED TRIALS						
No Comm teams	44.59%	*0.06*	54.40%	*0.06*	81.82%	*0.06*
Comm teams	32.68%	*0.08*	39.63%	*0.08*	82.14%	*0.06*

(B)

SET SIZE 48	Large change	*SE*	Small change	*SE*	No change	*SE*
RESPONSE TIME(in ms)						
Individuals	7043.72	*1576.92*	8996.83	*1512.61*	14942.06	*1996.52*
No Comm teams	4684.43	*524.40*	3969.79	*429.75*	11054.93	*1503.09*
Comm teams	3590.07	*403.10*	3717.43	*427.15*	10015.07	*877.63*
ACCURACY(% correct)						
Individuals	85.19%	*0.08*	55.56%	*0.10*	100%	*0.00*
No Comm teams	88.10%	*0.06*	77.38%	*0.08*	98.81%	*0.01*
Comm teams	95.24%	*0.02*	90.48%	*0.03*	100%	*0.00*
NUMBER OF FIXATIONS						
Individuals	18.44	*4.01*	11.38	*2.01*	39.41	*7.06*
No Comm teams	21.93	*2.22*	18.64	*2.20*	61.18	*9.77*
Comm teams	20.65	*1.50*	22.83	*3.22*	54.55	*6.58*
% RE-SCANNED TRIALS						
Individuals	97.22%	*0.02*	97.98%	*0.02*	100.00%	*0.00*
No Comm teams	90.81%	*0.02*	95.73%	*0.02*	100.00%	*0.00*
Comm teams	90.69%	*0.03*	92.77%	*0.02*	99.40%	*0.01*
% CROSS-SCANNED TRIALS						
No Comm teams	60.85%	*0.05*	53.59%	*0.08*	100.00%	*0.00*
Comm teams	11.06%	*0.03*	16.36%	*0.06*	71.43%	*0.10*

Fixations

No comm/Comm. Overall, fewer fixations were required with set size *6* than *48*, $F(1,6) = 57.87$, $p<.05$. However, averaging across set size and group conditions, the number of fixations varied across change magnitude, $F(2,12) = 28.07$, $p<.05$. Post-hoc comparisons showed more fixations were required on *no* compared to *small* and *large change* trials, $p<.017$. A change magnitude x set size interaction emerged, $F(2,12) = 22.82$, $p<.05$. Figure 9a, shows that the effect of change magnitude on fixations was driven by fixation variations with change magnitude when set size was *48*. A set size x change magnitude x group interaction also emerged, $F(2,12) = 7.33$, $p<.05$. Figure 10 shows that more fixations were required on *no change* trials, particularly with set size *48*, with a greater effect on *No comm* than *Comm dyads.*

Individuals/No comm. Overall, fewer fixations were required with set size *6* than *48*, $F(1,14) = 78.55$, $p<.05$. A set size x group interaction emerged, $F(1,14) = 4.66$, $p<.05$. Figure 11 shows that the effect of set size on number of fixations was due mostly to *No comm dyads* than *Individuals*. Averaging across set size and group, the number of fixations varied with change magnitude $F(2,28) = 49.87$, $p<.05$. Post-hoc comparisons showed more fixations on *no* compared to *small* and *large change* trials, $p<.017$. A set size x change magnitude interaction emerged, $F(2,28) = 38.34$, $p<.05$. Figure 9b shows that the change magnitude effect on fixations was driven by fixation variations with change magnitude when set size was *48*. Additionally, a set size x change magnitude x group interaction is depicted in Figure 12, $F(2,28) = 3.54$, $p<.05$. Both groups required more fixations on *no change* trials, particularly with set size *48*, with a greater effect on *No comm dyads* than *Individuals.*

Individuals/Comm. Overall, fewer fixations were required with set size *6* than *48*, $F(1,14) = 85.40$, $p<.05$. However, averaging across set size and group, the number of fixations required varied with change magnitude, $F(2,28) = 56.62$, $p<.05$, with more fixations on *no* compared to *small* and *large change* trials, $p<.017$. A set size x change magnitude interaction emerged, $F(2,28) = 28.31$, $p<.05$. Figure 9c shows that the change magnitude effect was driven by change magnitude effects on number of fixations when set size was *48*.

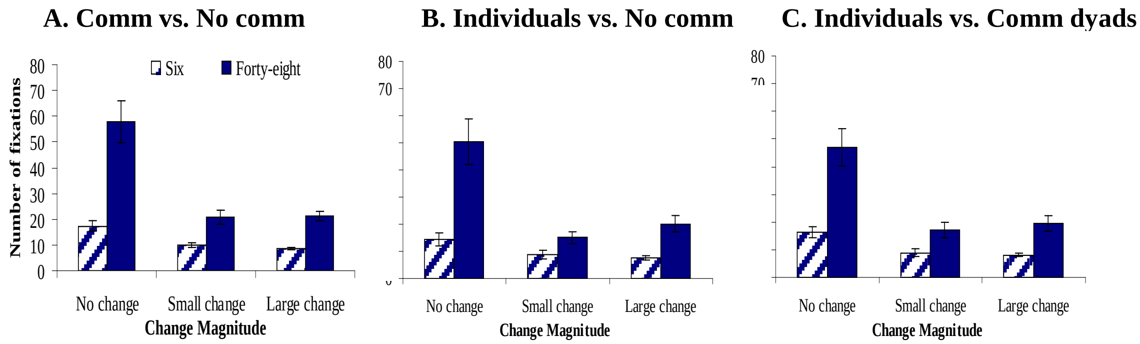


Figure 9. Number of fixations as a function of change magnitude for for set size 6 and 48 (set size x change magnitude interaction, p<.05). Each graph represents results from ANOVAs with the between-subjects factor indicated above the graph. Striped bars represent set size 6 and solid bars represent set size 48.

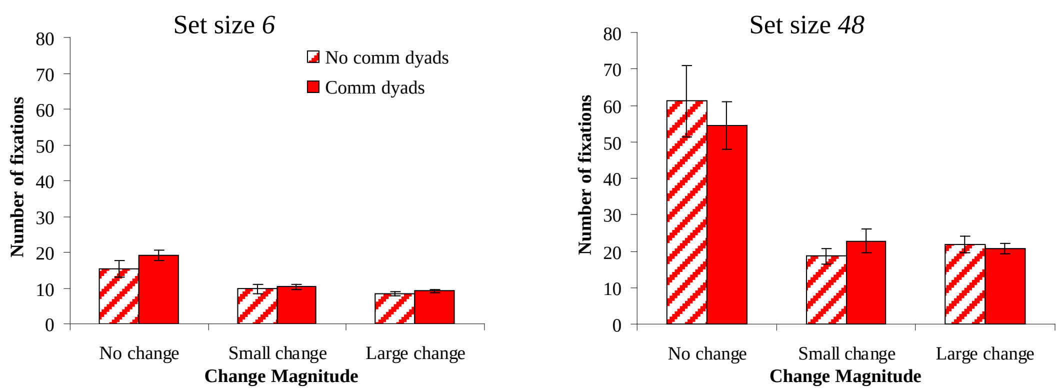


Figure 10. Number of fixations as a function of change magnitude for No comm and Comm dyads for set size 6 (on the left) and 48 (set size x change magnitude x group interaction, p<.05). Striped bars represent No comm dyads and solid bars represent Comm dyads.

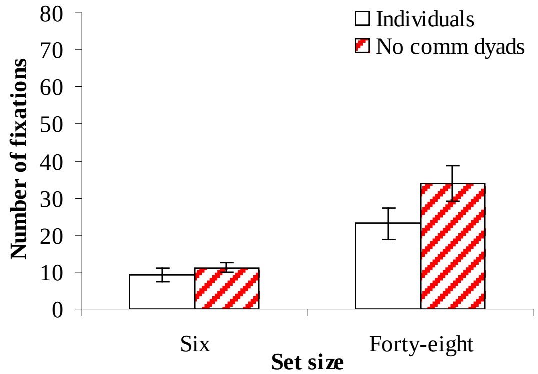


Figure 11. Number of fixations as a function of set size for Individuals and No comm dyads (set size x group interaction, p<.05). Unfilled bars represent Individuals and striped bars represent No comm dyads.

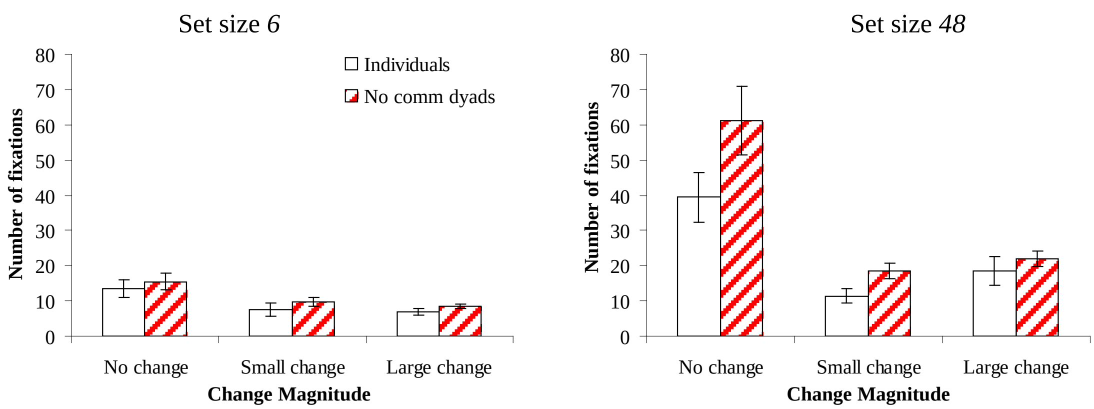


Figure 12. Number of fixations as a function of change magnitude for Individuals and No comm dyads for set size 6 (on the left) and set size 48 (set size x change magnitude x group interaction, p<.05). Unfilled bars represent Individuals and striped bars represent No comm dyads.

Cross-scanning

No Comm/Comm. Overall, *No comm dyads* cross-scanned more than *Comm dyads* $F(1,6) = 10.13$, $p<.05$, particularly with set size *48*, $F(1,6) = 6.86$, $p<.05$ (Figure 13; Group x set size interaction). Figure 14 depicts examples of typical trials for *No comm* and *Comm dyads*. Averaging across set size and group, percent of cross-scanned trials varied with change magnitude, $F(2,12) = 43.99$, $p<.05$, with more *no change* trials cross-scanned than *small*, or *large change* trials, $p<.05$. A change magnitude x group interaction emerged, $F(1,6) = 5.71$, $p<.05$. Figure 15 shows that the group effect on percentage of cross-scanned trials was greatest on *small change* trials and least on *no change* trials.

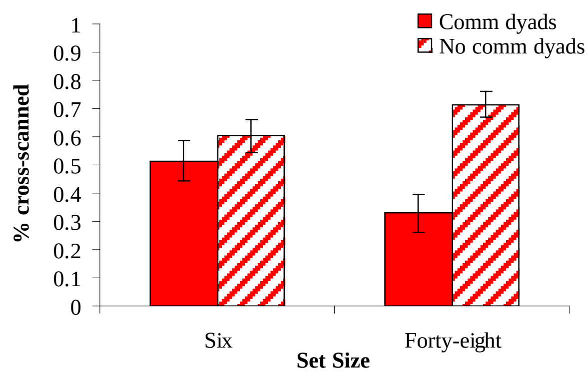


Figure 13. Percentage of cross-scanned trials as a function of set size for No comm and Comm dyads (set size x group interaction, p<.05). Solid bars represent Comm dyads and striped bars represent No comm dyads.

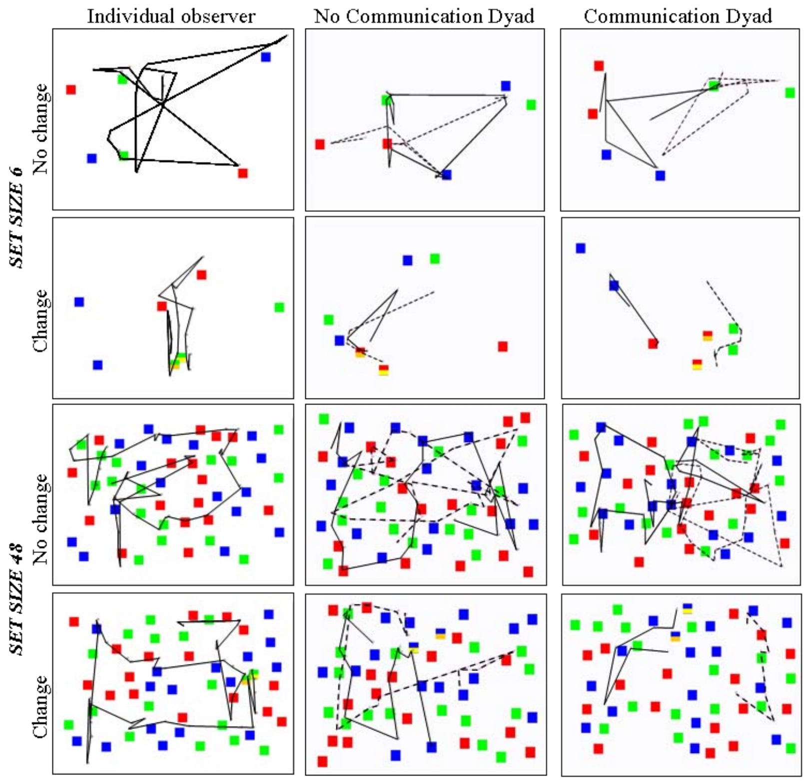


Figure 14. Scan paths for Individuals, No comm and Comm dyads for no change and change trials. The two-toned square is the square which changed in position. Solid lines in the Dyad trials represent the scan path of the observer that responded.

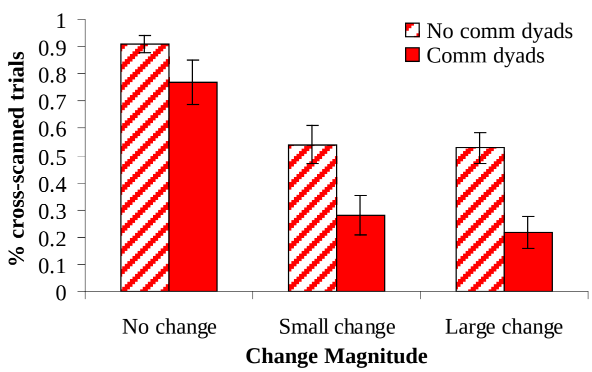


Figure 15. Percentage of cross-scanned trials as a function of change magnitude for No comm and Comm dyads (change magnitude x group interaction, p<.05). Striped bars represent No comm dyads and solid bars represent Comm dyads.

Re-scanning

No comm/Comm, Overall, more trials were re-scanned with set size *48* than *6*, $F(1,6) = 35.18$, $p<.05$. However, averaging across set size and group, percent of re-scanned trials varied across change magnitude, $F(2,12) = 43.10$, p<.05. Post-hoc comparisons showed that more *no change* trials were re-scanned compared to *small* and *large change* trials, $p<.017$. A set size x change magnitude interaction emerged, $F(2,42) = 18.35$, $p<.05$. Figure 16a shows that the change magnitude effect on percent of re-scanned trials was driven by greater variation in re-scanning when set size was *48*. Additionally, Figure 17 depicts a set size x group interaction, $F(1,13) = 6.82$, $p<.05$: *No comm dyads* re-scanned more trials with set size *48* than *6*, but *Comm dyads* re-scanned more trials when set size was *6* than *48.*

Individuals/No comm. A main effect of group was only marginally significant. On average, *Individuals* re-scanned on more trials than *No comm dyads*, $F(1,21) = 4.16$, $p = .05$. Overall, more trials were re-scanned with set size *48* than *6*, $F(1,21) = 32.46$, $p<.05$, and the percentage of re-scanned trials varied across change magnitude, $F(2,42) = 17.69$, $p<.05$. Post-hoc comparisons showed that more *no change* trials were re-scanned compared to *small* and *large change* trials, $p<.017$. A change magnitude x set size, $F(2,42) = 6.43$, $p<.05$ and change magnitude x group interaction emerged, $F(2,42) = 6.43$, $p<.05$. As shown in Figure 16b, the effect of change magnitude on percent of re-scanned trials was driven by variations in percent of re-scanned trials with change magnitude when set size was *48*, and also, as shown in Figure 18, by greater variations in re-scanning across change magnitude in *No comm dyads* than *Individuals*.

Individuals/Comm. On average, more trials were re-scanned when set size was *48* compared to *6*, $F(1,21) = 25.88$, $p<.05$, and percent of re-scanned trials varied by change magnitude, $F(2,42) = 14.03$, $p<.05$. Post-hoc comparisons showed that more *no change* trials were re-scanned than *small* or *large change* trials, $p<.017$. Additionally, a change magnitude x group interaction emerged, $F(2,42) = 4.78$, $p<.05$. Figure 19 shows that variations in percentage of re-scanned trials across change magnitude was greater in *Comm dyads* than *Individuals.*

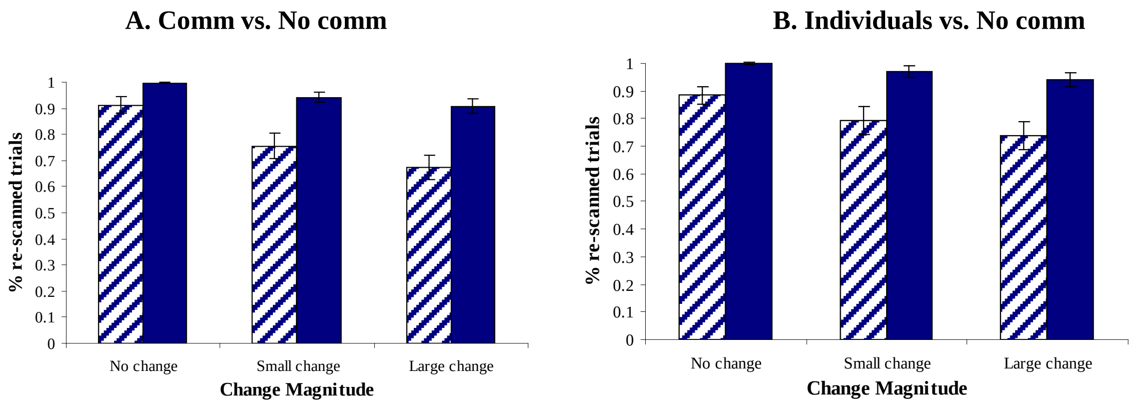


Figure 16. Percentage of re-scanned trials as a function of change magnitude for set size 6 and 48 (set size x change magnitude interaction, p<.05). Each panel represents results from ANOVAs with the between-subjects factor indicated above the graphs. Striped bars represent set size 6 and solid bars represent set size 48.

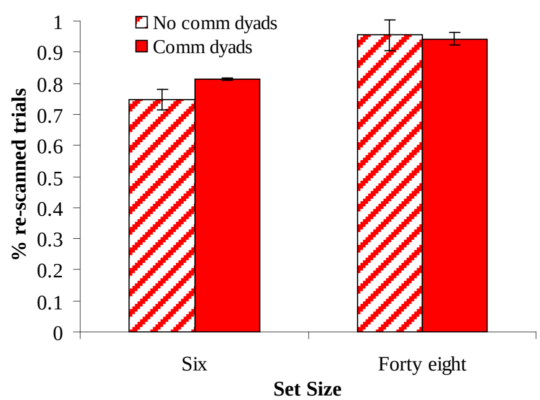


Figure 17. Percentage of re-scanned trials as a function of set size for No comm and Comm dyads (set size x group interaction, p<.05). Striped bars represent No comm dyads and solid bars represent Comm dyads.

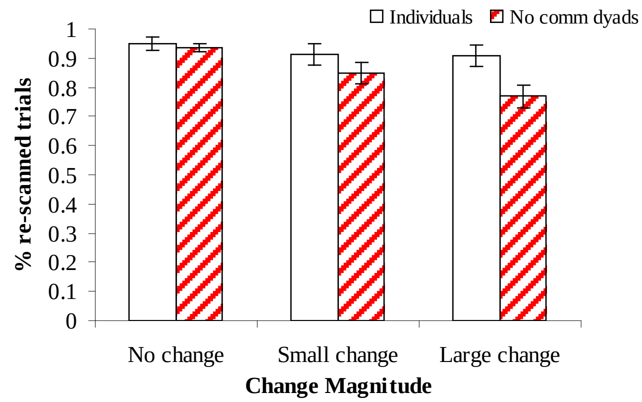


Figure 18. Percentage of re-scanned trials as a function of change magnitude, averaged across set size (change magnitude x group interaction, p<.05). Unfilled bars represent Individuals and striped bars represent No comm dyads.

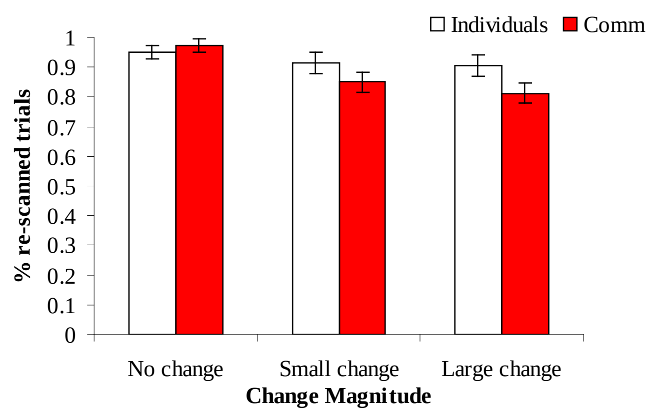


Figure 19. Percentage of re-scanned trial as a function of change magnitude for Individuals and Comm dyads (change magnitude x group interaction, p<.05). Unfilled bars represent Individuals and solid bars represent Comm dyads.

Workload: NASA RTLX

The set size manipulation hypothesized to drive workload was validated by the data. As indicated below, all ANOVAs with RTLX yielded significant main effects of set size. See Table 3 below for RTLX means and standard errors for all groups.

No comm/Comm, Overall, *No comm dyads* (M = 42.71, SE = 6.13) had higher RTLX scores than *Comm dyads* (M = 35.54, SE = 5.78), $F(1,6) = 17.15$, $p<.05$. Additionally, averaging across group, ratings varied by set size, $F(1,6) = 13.56$, $p<.05$.

Individuals/No comm. Only a main effect of set size emerged from this analysis. On average, both *Individuals* and *No comm dyads* gave higher workload ratings when set size was *48* (M = 45.54, SE = 6.84) compared to *6* (M = 35.20, SE = 5.70), $F(1,14) = 10.23$, $p < .05$.

Individuals/Comm. Only a main effect of set size emerged from this analysis. Overall, both *Individuals* and *Comm dyads* gave higher workload ratings when set size was *48* (M = 42.82, SE = 6.80) compared to *6* (M = 30.74, SE = 5.38), $F(1,14) = 12.89$, $p < .05$.

Table 3. Means and SEs for RTLX, heart rate, and NNI for Individuals, No comm and Comm dyads and Team workload scale (TWS) for Comm dyads.

	SET SIZE 6	*SE*	SET SIZE 48	*SE*
RTLX				
Individuals	30.54	*2.03*	45.50	*5.30*
No Comm dyads	39.86	*1.78*	45.57	*1.47*
Comm dyads	30.95	*1.51*	40.14	*1.85*
Heart rate				
Individuals	907.37	*67.88*	874.24	*49.49*
No Comm dyads	862.12	*15.15*	844.49	*10.69*
Comm dyads	847.23	*11.14*	835.33	*9.31*
NNI				
Individuals	0.86	*0.02*	0.87	*0.01*
No Comm dyads	0.84	*0.02*	0.85	*0.01*
Comm dyads	0.85	*0.02*	0.86	*0.02*
TWS				
Comm dyads	3.40	*0.79*	3.87	*0.55*

Workload: NASA TLX subscale ratings

No comm/Comm, Overall, *No comm dyads* (M = 42.71, SE = 6.13) gave higher ratings than *Comm dyads* (M = 35.54, SE = 5.78), $F(1,6) = 17.15$, $p < .05$. Additionally, averaging across group, ratings varied by set size with higher ratings for set size *6* (M = 35.40, *S.E.* = 6.17) than set size *48 6* (M = 42.86, *S.E.* = 5.74), $F(1,6) = 13.56$, $p < .05$, and by subscale, $F(1,6) = 30.18$, $p < .05$, with the highest ratings for performance (M= 81.12, *S.E.* = 4.46), followed by mental demand (M = 40.68, *S.E.* = 8.01), then by effort (M = 37.66, *S.E.* = 6.71), then by temporal demand (M = 35.27, *S.E.* = 6.81), then by frustration (M = 22.98, *S.E.* = 6.08), and last by physical demand (M = 17.07, *S.E.* = 3.67). A set size x subscale interaction is depicted in Figure 20. The effect of subscale on ratings appears to be driven by greater

variability in ratings across subscale for set size *48* compared to set size *6*, $F(5,30) = 3.94$, $p<.05$. Additionally, a group x subscale interaction also emerged, $F(5,30) = 8.21$, $p<.05$. As shown in Figure 21, compared to *No Comm dyads, Comm dyad* gave lower ratings on all subscales except for the performance scale, in which they rated their performance higher than *No Comm dyads*.

Individuals/No comm. Overall, ratings were higher set size was *48* ($M = 45.54$, $SE = 6.84$) compared to *6* ($M = 35.20$, $SE = 5.70$), $F(1,14) = 10.23$, $p<.05$. Additionally, ratings varied by subscale, $F(5,70) = 32.06$, $p<.05$, with highest ratings for performance ($M= 76.39$, $S.E. = 6.60$), followed by mental demand ($M = 46.85$, $S.E. = 7.12$), then by effort ($M = 43.50$, $S.E. = 6.59$), then by temporal demand ($M = 37.38$, $S.E. = 7.17$), then by frustration ($M = 22.85$, $S.E. = 6.50$), and last by physical demand ($M = 14.23$, $S.E. = 3.61$). A set size x subscale interaction is depicted in Figure 22. The effect of subscale on ratings appears to be driven by greater variability in ratings across subscale for set size *48* compared to set size *6*, $F(5,70) = 9.41$, $p<.05$.

Individuals/Comm. Overall, both *Individuals* and *Comm dyads* gave higher workload ratings when set size was *48* ($M = 42.82$, $SE = 6.80$) compared to *6* ($M = 30.74$, $SE = 5.38$), $F(1,14) = 12.89$, $p<.05$. Additionally, ratings varied by subscale, $F(5,70) = 40.92$, $p<.05$, with highest ratings for performance ($M= 79.12$, $S.E. = 6.77$), followed by mental demand ($M = 40.60$, $S.E. = 7.56$), then by effort ($M = 37.44$, $S.E. = 7.01$), then by temporal demand ($M = 30.22$, $S.E. = 7.22$), then by frustration ($M = 19.30$, $S.E. = 5.26$), and last by physical demand ($M = 14.01$, $S.E. = 3.34$). A set size x subscale interaction is depicted in Figure 23. The effect of subscale on ratings appears to be driven by greater variability in ratings across subscale for set size *48* compared to set size *6*, $F(5,70) = 6.41$, $p<.05$.

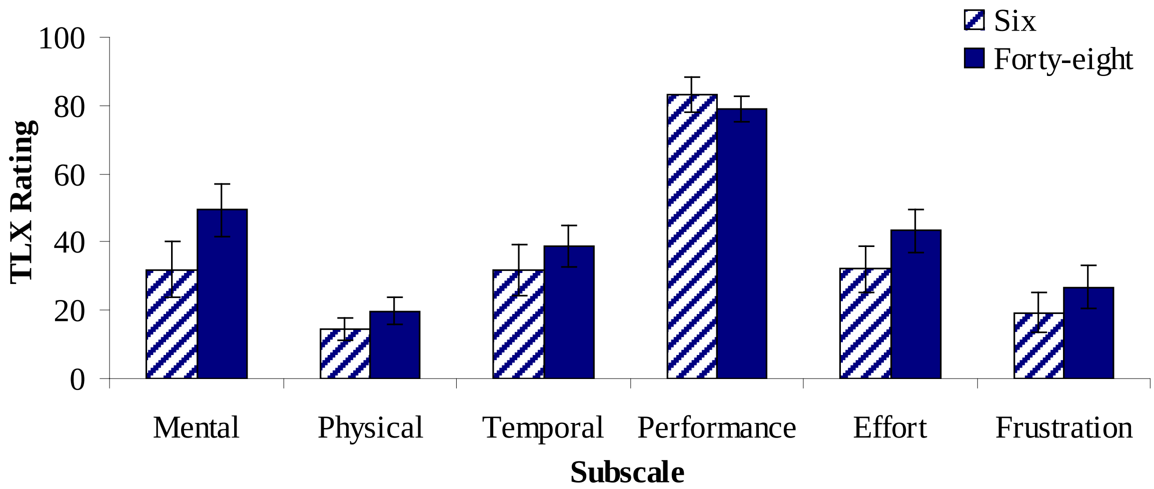


Figure 20. TLX ratings as a function of subscale for set size 6 and 48, averaged across Comm and No comm dyads (set size x subscale interaction, p<.05).

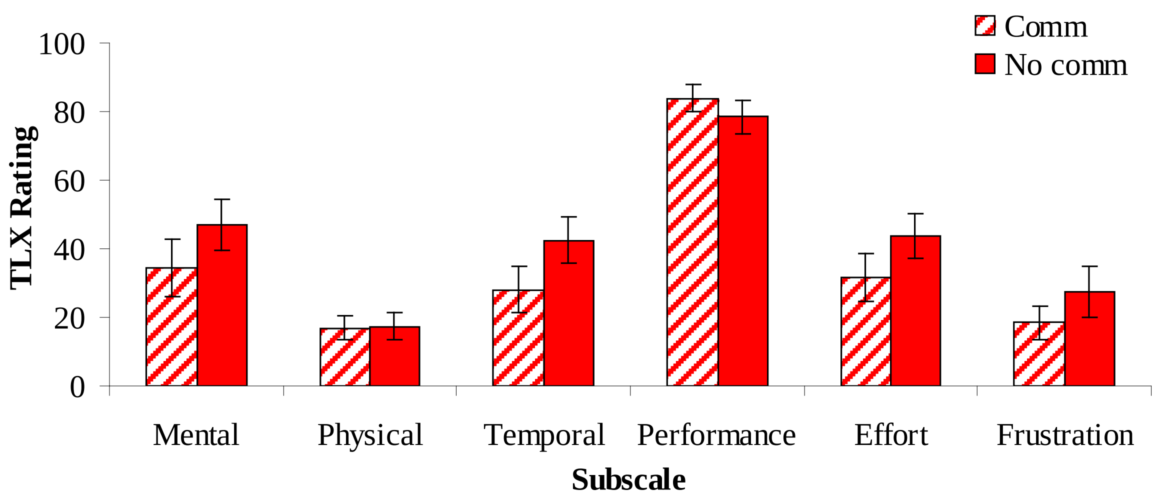


Figure 21. TLX ratings as a function of subscale for Comm and No comm dyads, averaged across set size (group x subscale interaction, p<.05).

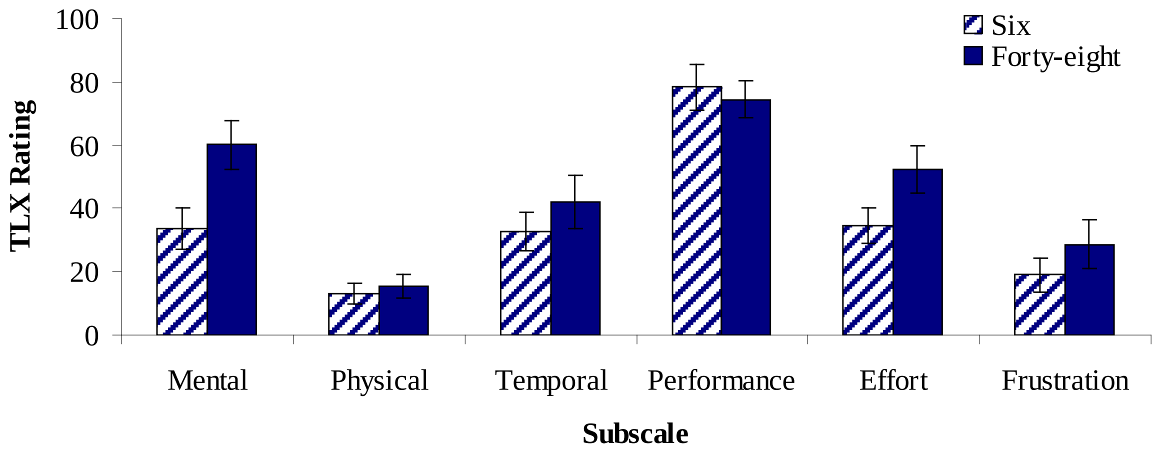


Figure 22. TLX rating as a function of subscale for set size 6 and 48, collapsed across Individuals and No comm dyads (group x subscale interaction, p<.05).

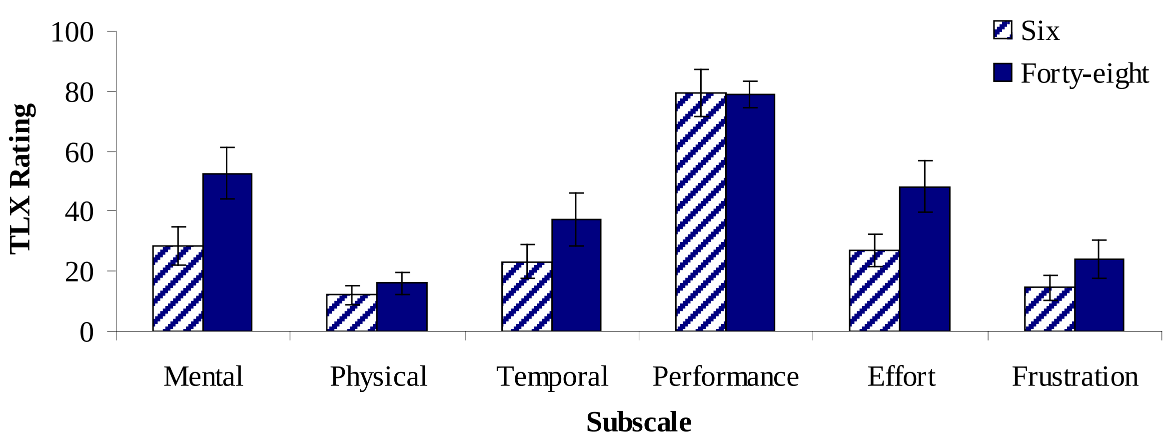


Figure 23. TLX rating as a function of subscale for set size 6 and 48, collapsed across Individuals and Comm dyads (group x subscale interaction, p<.05).

Workload: Heart Rate

ANOVAs yielded no main effects or interactions with group, and only marginally significant main effects of set size. Therefore, data from all groups were combined to increase *n* and compared in a paired t-test with set size as the independent variable $t(18) = 2.14$, $p<.05$. Heart rate was slower when set size was *6* (M = 868.54, SE = 24.61) compared to *48* (M = 848.94, SE = 18.15), suggesting lower workload, overall, with set size *6* than *48*.

Workload: NNI

No comm/Comm, Overall, *No comm dyads* (M = .86, SE = .01) yielded higher NNIs, thus more random fixation patterns, than *Comm dyads* (M = .82, SE = .01), $F(1,13) = 10.07$, $p<.05$. A group x set size interaction emerged, $F(1,13) = 6.15$, $p<.05$. As shown in Figure 24, the effect of group on NNI was driven by differences when set size was *48*, such that *No comm dyads* yielded particularly higher NNIs than *Comm dyads* when set size was *48*.

Individuals/No comm. No significant main effects or interactions emerged from this comparison.

Individuals/Comm. The main effect of group and the interaction of group with set size were marginally significant. Individuals had higher NNIs (M = .86, SE = .01), than *Comm dyads* (M = .82, SE = .01), $F(1,14) = 4.43$, $p = .05$. A group x set size interaction was also marginally significant, $F(1,21) = 4.23$, $p = .06$. Figure 25 shows that *Individuals* yielded particularly higher NNIs than *Comm dyads* when set size was *48*.

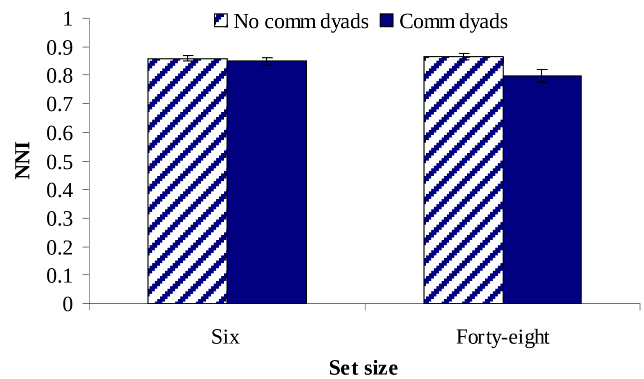


Figure 24. NNI as a function of set size 6 and 48 for No comm and Comm dyads (group x set size interaction, p<.05). Striped bars represent No comm dyads and solid bars represent Comm dyads.

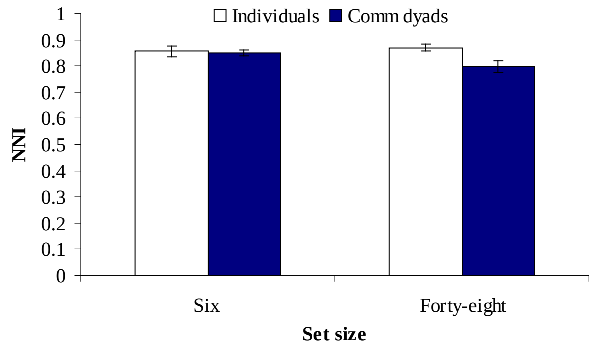


Figure 25. NNI as a function of set size 6 and 48 for Individuals and Comm dyads (group x set size interaction, p<.05). Unfilled bars represent No comm dyads and solid bars represent Comm dyads.

Workload: Team Workload Scale

Overall, average team workload rating varied by subscale, $F(4,24) = 6.10$, $p<.05$. Communication demands were rated highest ($M = 4.60$, $SE = .75$), followed by Leadership demands ($M = 4.12$, $SE = .77$), Mental demands ($M = 4.10$, $SE = .71$), Coordination demands ($M = 3.20$, $SE = .59$), with Control demands rated lowest ($M = 2.20$, $SE = .53$). (A set size x subscale interaction emerged, $F(4,24) = 3.67$, $p<.05$. Figure 26 shows that variability in average rating across subscale was driven by greater variability when set size was *48* than *6*. Post-hoc pairwise comparisons of subscales by set size did not yield any significant differences after Bonferroni adjustment to alpha (alpha level .05/5 = .01). Only the *Communication* subscale was marginally significant before Bonferroni correction was applied, $t(6) = -2.309$, $p = .06$. As shown in Figure 26, average rating for *Communication demands* were higher when set size was *48* ($M = 5.43$, $SE = .83$) compared to *6* ($M = 3.76$, $SE = .67$).

Confidence Ratings

An average confidence rating was obtained from the two members in a *Dyad*. Only 7 of the 9 *Individuals* completed exit questionnaires. Therefore, a mean score was obtained with confidence ratings from the 7 participants in the *Individual* condition and used to fill in the missing scores for this analysis. A paired t-test was used to compare *Comm* and *No*

comm dyads, and unpaired t-tests were used to compare *Individuals* to *Comm dyads* and *Individuals* to *No comm dyads*. Only the *No comm/Comm dyad* comparison yielded statistically significant results. *Comm dyads* were more confident in their responses (M = 9.07, *S.E.* = .35) compared to *No comm dyads*, $t(6) = 3.06$, $p<.05$ (M = 8.43, *S.E.* = .67). The comparison of *Individuals* (M = 8.29, *S.E.* = 1.09) to *Comm dyads* (M = 9.07, *S.E.* = .35) was only marginally significant, $t(14) = 1.83$, $p = .089$ with *Comm dyads* more confident in their responses than *Individuals*.

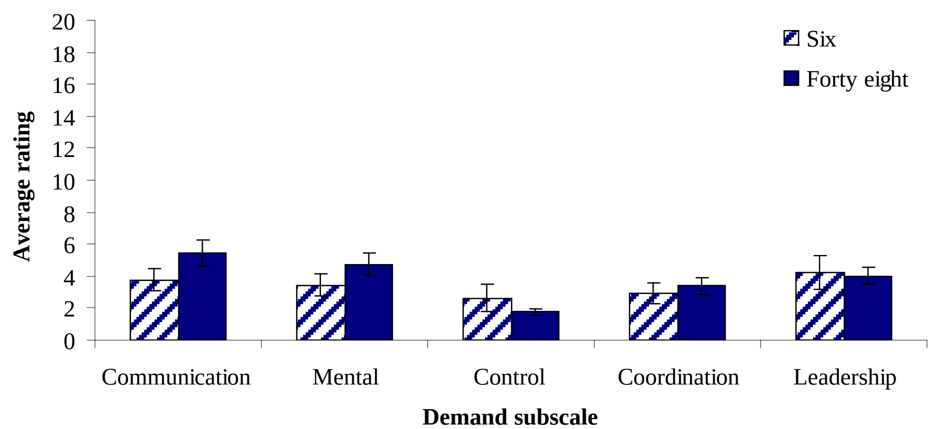


Figure 26. Average dyad rating as a function of team workload demand subscale for set size 6 and 48 (set size x subscale interaction, p<.05). Striped bars represent set size 6 and solid bars represent set size 48.

NNI as a Workload Measure

To perform the correlations and regressions between workload measures, data from dyads which were averaged together for the group comparisons were again separated by dyad member. Dyad members were treated as individual participants.

Evaluation of the effectiveness of NNI as a workload measure proceeded in the following manner: First, RTLX was correlated to set size to confirm that the set size manipulation hypothesized to affect workload was reflected in the individual RTLX scores. Although the measure was validated by main effects of set size from the ANOVAs performed on the data, this analysis was performed to confirm the sensitivity of the RTLX with the use of the independent dyad member RTLX scores. Second, NNI, heart rate and RTLX were correlated following confirmation of sensitivity to set size differences of individual dyad member RTLX scores. However, only a subset of the data from set size *48* were included in the correlations of workload measures due to the following: (1) ceiling effects on NNIs from trials with set size *6* (see Figures 24, 25 and section below), and (2) heart rate data were incomplete as heart rate was not monitored for 4 participants in the *Individual* observer condition due to technical difficulties with the heart rate monitor.

Ceiling effect and confounding factor on NNI analysis

NNI is the ratio of the average distance between observed and randomly distributed points in a specified area. Our NNI data from the set size *6* condition yielded NNIs close to 1 indicating relatively random fixation patterns. More importantly, our data suggest higher NNIs, hence higher workload, with set size *6* compared to set size *48*. This is in contrast to previous findings by Di Nocera, et al. (2006; 2007) and both our performance and NASA TLX data which indicate that change detection was more difficult with set size *48* than with set size *6.*

Due to inconsistency with both performance and subjective workload data, we determined that the ceiling effect on the NNI data from set size *6* were likely driven by a confounding factor inherent in the design of the experiment - icons were randomly positioned in every trial to counteract potential learning effects. As a result, the random distribution of icons in each display was reflected by random fixation distributions and NNIs close to 1. Although icons were randomly positioned in both set size *6* and set size *48*, some dyads in

the *Comm* condition, in addition to all the dyads in the *No comm* condition, searched the entire screen independently for the changing icon only when set size was *6*. In contrast, in the set size *48* condition, all dyads in the *Comm* condition divided the screen so that each member searched half of the screen for a change in icon position. In some cases, one member would ask the other to confirm that none of the icons were changing in position which resulted in cross-scanning. However, as reflected in the scan paths of *Comm dyads* with set size *48* (Figure 14), dyads divided the search space in half on most trials. This effectively reduced the variability in distances between fixation points by reducing the space in which more points, relative to set size *6*, were distributed. Thereby, on average, the likelihood that fixations would approach a random distribution or an NNI value close to 1 was lower when set size was *48* compared to when set size was *6*.

In summary, the confounding factor described above was due to the difference in search space across task difficulty or set size conditions. While icons in each display were randomly distributed across both set sizes, set size *6* had fewer icons and a wider area in which icons were distributed, and thereby greater potential variability in distances between points than with set size *48* trials. As a consequence, the likelihood that a distribution of fixations would be close to a random distribution was greater, on average, in set size *6* than set size *48*. More importantly, compared to set size *48*, the displays in set size *6* were easier to search through without a strategy since there were few icons within the field of view. Although a strategic search for the changing icon may have benefited performance with both set sizes, a random search was effective only when few icons were on the display, such as on set size *6* trials.

RTLX and set size

Results indicate a positive correlation between RTLX and set size, with RTLX higher for the larger set size, set size *48*, $R^2 = .12$. $t = 3.10$, $p<.05$.

NNI and RTLX

NNI and RTLX are not significantly correlated.

NNI, heart rate, and RTLX

Correlations between NNI and heart rate, NNI and RTLX, and heart rate and RTLX were not statistically significant.

Regression model of change detection performance with multiple observers

Data from *Comm* and *No comm dyads* were submitted to a backward step-wise regression analysis. Our goal in building the models was to identify the best predictors of efficient change detection with multiple observers. Such a model could inform both the development of training methods for change detection and adaptive technologies to support team performance.

RT and accuracy were used as dependent variables in 2 independent models of change detection performance. Predictors were number fixations, percentage of re-scanned and cross-scanned trials, heart rate, and RTLX. NNI was not included in the analysis due to the confounding factor on data from set size *6*, as discussed above.

Response time model

Number of fixations and RTLX were the only significant predictors of RT, $F(2,25) = 63.64$, $p<.05$, see Tables 3 for the ANOVA and regression tables. The regression equation $y' = -660.04 + 162.441\text{Fixations} + 27.421\text{RTLX}$ predicts 82.3% of the variance in RT in change detection with multiple observers, adjusted $R^2 = .823$. See Table 3a and 3b for the regression and ANOVA tables.

Table 3. Regression and ANOVA tables for the model of RT in change detection with multiple observers.

(A)

Effect	Coefficient	Std Error	Std Coef	Tolerance	t	P(2 Tail)
CONSTANT	-660.04	689.33	0	.	-0.958	0.347
FIXATIONS	162.441	15.133	0.881	0.975	10.734	0.000
RTLX	27.421	15.96	0.141	0.975	1.718	0.098

(B)

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
Regression	1.27E+08	2	6.35E+07	63.638	0.000
Residual	2.49E+07	25	997304.883		

Accuracy model

All but one entered into the model were significant predictors of accuracy, $F(5,22) = 15.21$, $p<.05$, see Tables 4 for the ANOVA and regression tables. RTLX reached marginal significance at $p = .067$. The regression equation y' = 2.192 + .002Fixations -.095Cross-scan % -.603Re-scan % +-.001RTLX + -.001heart rate and predicts 72.5% of the variance in accuracy of change detection with multiple observers, adjusted $R^2 = .732$. See Table 4 for the regression and ANOVA tables.

Table 4. Regression and ANOVA tables for the model of accuracy in change detection with multiple observers.

(A)

Effect	Coefficient	Std Error	Std Coef	Tolerance	t	P(2 Tail)
CONSTANT	2.192	0.175	0.000	.	12.498	0.000
FIXATIONS	0.002	0.001	0.409	0.381	2.500	0.020
CROSS-SCAN	-0.095	0.039	-0.254	0.957	-2.465	0.022
RE-SCAN	-0.603	0.129	-0.855	0.307	-4.690	0.000
RTLX	-0.001	0.001	-0.226	0.738	-1.924	0.067
HEART RATE	-0.001	0.000	-0.796	0.548	-5.838	0.000

(B)

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
Regression	0.127	5.000	0.025	15.207	0.000
Residual	0.037	22.000	0.002		

Discussion

The advantage provided by having more eyes/resources for change detection depends on whether or not (1) observers are allowed to communicate, and (2) a change occurs. Our results provide direct evidence that communication allows efficient attention allocation and scanning strategies for change detection by allowing observers to function as a team. The results of the regression analyses predicting performance solidify the importance of attention allocation and scanning strategies in change detection with multiple observers.

Overall, all workload measures used in the present study, including NNI, were sensitive to changes in task difficulty. Additionally, RTLX and NNI both reflected differences in workload between communicating and non-communicating dyads, with both measures indicating higher workload in non-communicating compared to communicating dyads. However, none of the workload measures were statistically correlated to one another and results of the NNI analysis revealed some limitations for using NNI to index workload.

Are effects of communication on change detection performance apparent in eye movements and scan paths?

Performance results are consistent with those from Tollner et al. (2006). Overall, communicating dyads detected changes faster and more accurately than individuals. Likewise, communicating dyads outperformed both non-communicating dyads and individual observers when the task was difficult. Our goal was to determine if the results of Tollner et al. (2006), replicated in the present study, may be attributed to attention allocation and scanning strategies afforded not just by availability of resources, but by the ability of multiple observers to communicate and function as a team.

First, our data show that communicating dyads functioned as a team, with several major components of teamwork evident in their scan paths: back-up behavior, coordination, feedback, and, of course, communication (Dickinson & McIntyre, 1997). Although cross-scanning was evident in both communicating and non-communicating dyads, cross-scanning in communicating dyads was less frequent and deliberate for confirmation of change or no change. Additionally, designated areas of responsibility are evident in the scan paths of

communicating dyads only (see Figure 14). In contrast, the inferior performance of individual observers and non-communicating dyads may be due to the fact that they served as their own feedback mechanism, and lacked the ability to coordinate their search strategies for optimal performance.

Second, analyses of both fixations and scan paths support the idea that individuals and teams utilize different attention allocation and scanning strategies in change detection. Non-communicating dyads (1) required more fixations, thereby more re-allocation of attention, than communicating dyads and individuals to detect changes, and (2) cross-scanned more trials than communicating dyads, particularly when the task was difficult (set size 48). However, non-communicating dyads gained no obvious performance benefits over individuals and communicating dyads. Instead, they were slower than individuals and less accurate than communicating dyads when the task was difficult. Overall, non-communicating dyads used more time-consuming, less efficient strategies than both individuals and communicating dyads. In addition, these findings suggest that the benefits of multiple observers in change detection are undermined by the lack of a medium such as communication for coordinating actions and determining strategies.

Results from our eye movement analyses are consistent with Zelinsky's (2001) findings that eye movements in change detection reflect information gathering and change confirmation. This is evident in the use of re-scanning by all three groups and cross-scanning in communicating dyads. More importantly, results showed that communication allowed dyads to develop strategies for efficient resource allocation in the service of information gathering and change confirmation. While cross-scanning appeared to be a costly, tacit behavior for non-communicating dyads, it served as an effective strategy for confirmation of change for communicating dyads.

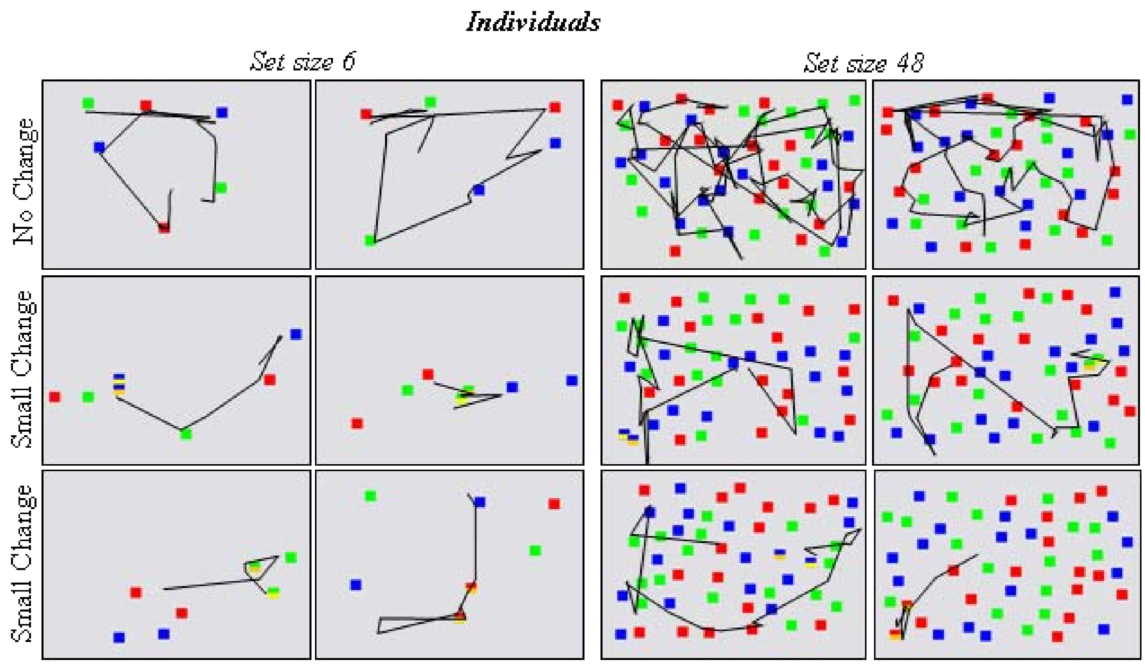
Interestingly, while communicating dyads were better than individuals at detecting changes, trials that lacked change were costly for both individuals and dyads regardless of communication condition. On *no change* trials, both individual observers and dyads were slower to respond and required more fixations. This finding has implications for environments such as tactical C2 in which a common and current understanding of a visual environment is necessary for decision-making. Delayed information updating due to slow confirmation of the lack of change by one or more members of a distributed team could delay

decision-making and potentially lead to catastrophic outcomes. Likewise, cognitive resources consumed by slow confirmation of no change may be better allocated to other important tasks.

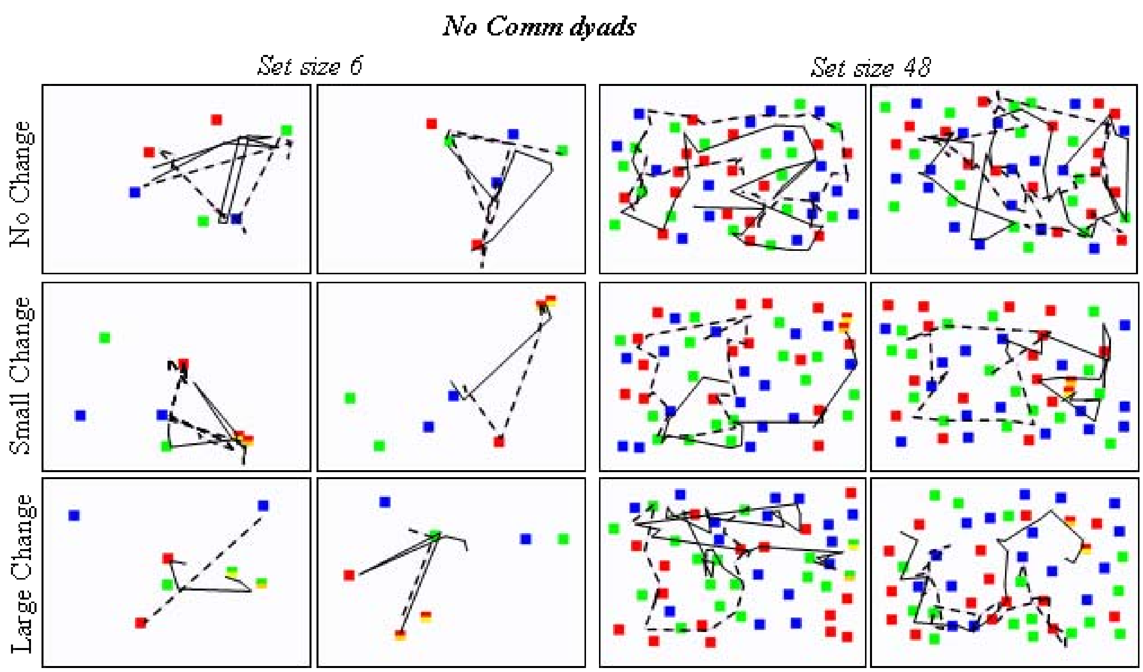
Finally, Tollner et al. (2006) speculated that the slowed responses on trials in which no change occurred may have resulted from a serial and exhaustive scanning of the entire display, as previously seen in visual search studies (e.g., Treisman & Gelade, 1980). Our fixation data provide some support for this speculation. On trials with no change, on average, individuals required approximately 13 fixations with set size *6* and 40 fixations with set size *48*. Likewise, communicating dyads required approximately, 19 fixations with set size *6* and 54 with set size *48*, non-communicating dyads required approximately 15 fixations with set size *6* and 61 with set size *48*. These fixation numbers are approximately twice those from trials in which a change occurred and are almost equal to the number of stimuli in the display. Moreover, visual inspection of scan paths suggest that more of the display was searched on a majority of trials when no change occurred compared to when a change did occur (see Figure 14 and Figure 27). However, a more thorough analysis of scan paths is necessary to confirm this hypothesis. Figure 27 below are representative samples of scan paths from individual observers, communicating and non-communicating dyads.

Findings on scanning strategies may be used for training and development of adaptive aids and technologies to support team performance. First, evidence suggests that scanning strategies used in search can be trained (Wang, Lin, & Drury, 1997). Therefore, identification of efficient scanning patterns may be used in training both individual operators and teams to search for relevant changes more effectively. Second, identification of inefficient scanning patterns may be applied to the development of adaptive aids for individuals and teams showing signs of inefficient change detection or search behavior. Finally, an understanding of when change detection strategies could delay decision-making can guide the design of interfaces used by highly synchronized, distributed teams, such as teams in tactical C2. The importance of understanding the patterns of eye movements for both training and development of adaptive aids is supported by our models which revealed that number of fixations, re-scanning, and cross-scanning are significant predictors of either change detection speed, or change detection accuracy, or both.

(A)


Individuals
Set size 6
Set size 48
No Change
Small Change
Small Change

(B)


No Comm dyads
Set size 6
Set size 48
No Change
Small Change
Large Change

(C)

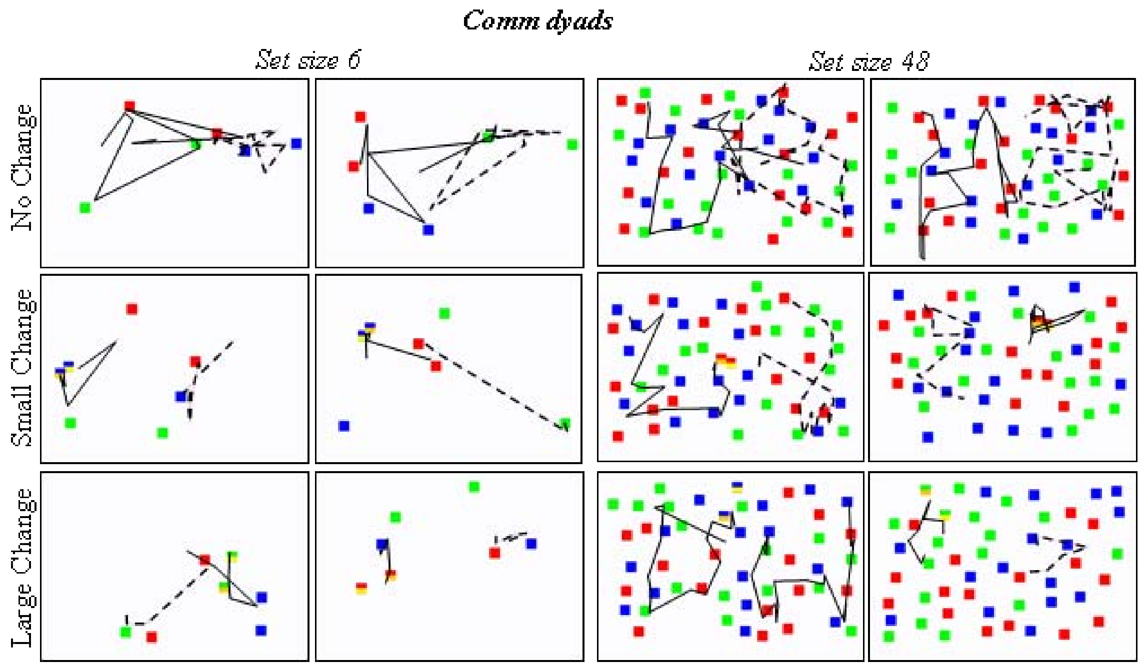


Figure 27. (A) Scan paths for Individuals and No comm dyads for trials with No change, Small and Large change. Scan paths of Individuals are shown against a gray background. (B) A greater portion of No change displays were scanned compared to Small and Large change displays. (C) Scan paths for Comm dyads for trials with No change, Small, and Large change.

Is the NNI a viable workload index?

Our data suggest that the NNI is sensitive to variations in task difficulty and subjective workload, but with certain constraints. First, given that the NNI is the ratio of the average distance between actual and random points within a space, it may be a valid workload metric only when the comparison is between conditions with the same relevant visual space. More importantly, this is regardless of whether or not the task within the space is static or dynamic in nature. The metric may fall short as a workload index in experiments such as the present in which the size of the relevant visual space is uncontrolled and may vary both by condition and by participant. Second, the reliability of the metric itself depends on the number of points used in the calculation with more points yielding more reliable the NNI. Therefore, the measure may be unreliable when used to measure workload in tasks that require or allow

few fixations, regardless of difficulty. For example, in air battle management, situations may vary within relatively short periods of time, thereby allowing only few fixations within each relevant time period. Lastly, the lack of significance in the correlation between NNI and other workload measures indicates that more work is necessary to validate the NNI as a reliable index of cognitive workload.

References

Ascension Technology Corporation. http://www.ascension-tech.com/products/flockofbirds.php

Byers, J.C., Bittner, Jr, A.C. & Hill, S.G. (1989). Traditional and Raw Task Load Index (TLX) Correlations: Are Paired Comparisons Necessary? In: Mital, A. (ed.): Advances in Industrial Ergonomics and Safety I, Taylor and Francis 481-485.

Carpenter, R.H.S. (1988). Movements of the eyes. (2nd ed.) London, Pion.

Castiello, U. & Umilta, C. (1990). Size of the attentional focus and efficiency of processing. *Acta Psychologica* (*Amst*), *73*, 195-209.

Dickinson, T.L. & McIntyre, R.M. (1997). A conceptual framework for teamwork measurement. In M. T. Brannick, E. Salas & C. Prince (Eds.), *Team performance assessment and measurement. Theory, methods, and applications*. (pp. 19-43). Mahwah, NJ: Lawrence Erlbaum.

Di Nocera, F., Camilli, M., & Terenzi, M. (2007). A random glance to the flight deck: pilot's scanning strategies and the real-time assessment of mental workload. *Journal of Cognitive Engineering and Decision Making.*

Di Nocera, F., Terenzi, M., & Camilli, M. (2006). Another look at scanpath: distance to nearest neighbour as a measure of mental workload. In D. de Waard, K.A. Brookhuis, and A. Toffetti (Eds.), *Developments in Human Factors in Transportation, Design, and Evaluation* (pp. 295-303). Maastricht, The Netherlands: Shaker Publishing.

Eriksen, C.W., & Yeh, Y.Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance, 11,* 583-597.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.) *Human MentalWorkload*. North-Holland: Elsevier Science Publishers (pp. 139-183).

Hildebrand, G. A., Pharmer, J. A., Weaver, M. D. (2003). Performance and usability evaluation of an automated task management concept prototype design. *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting*, 1913-1917.

Huynh, H., & Feldt, L.S. (1970). Conditions under which mean square ratios in repeated measures designs have exact f-distributions. *Journal of the American Statistical Association, 65*, 1582-1589.

McCarley, J.S., Vais, M., Pringle, H., Kramer, A.F., Irwin, D.E., & Strayer, D.L. (2004). Conversation disrupts scanning and change detection in complex visual scenes. *Human Factors*, *46*, 424-436.

Muthard, E.K., & Wickens, C.D. (2002). *Factors that mediate flight plan monitoring and errors in plan revision: An examination of planning under automated conditions* (AHFD-02-11/NASA-02-8). Savoy, IL: University of Illinois, Aviation Human Factors Division.

Posner, M.I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32,* 3-25.

Rensink, R.A., O'Regan, J.K., & Clark, J.J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8,* 368-373.

Shepherd, M., Findlay, J.M. & Hockey, G.R.J. (1986). The relationship between eye movements and spatial attention. *Quarterly Journal of Experimental Psychology, 38A*, 475-491.

Simons, D.J., & Levin, D.T. (1997). Change blindness. *Trends in Cognitive Science*, *1,* 261-267.

Tollner, A.M., Riley, M.A., Nelson, W.T., Shockley, K.D.Cummins-Sebree, S. (2006). Individual and Team Susceptibility to Change Blindness. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*, San Francisco, CA.

Treisman, A.M. & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12,* 97-136.

Wang, M., Lin, S., & Drury, C. (1997). Training for strategy in visual search. *International Journal of Industrial Ergonomics, 20*, 101-108.

Zelinsky. G.A. (2001). Eye movements during change detection: Implications for search constraints, memory limitations, and scanning strategies. *Perception & Psychophysics*, 63(2), 209-225.

Appendix A: Consent Form and Biographical Questionnaire

INFORMATION PROTECTED BY THE PRIVACY ACT OF 1974

Informed Consent Document

For

Individual and Team Susceptibility to Change Blindness

AFRL/HECP, WPAFB, OH, Building 33

Principal Investigator: W. Todd Nelson, DSN-785-8803, AFRL/HECP, william.nelson@wpafb.af.mil

Associate Investigators: Rebecca D. Brown, DSN-785-0884, AFRL/HECP, becky.brown@wpafb.af.mil

1. **Nature and purpose:** You have been offered the opportunity to participate in the "Individual and Team Susceptibility to Change Blindness" research study. Your participation will occur sometime between 01 November 2006 and 01 November 2007, in the BCM2 Lab at Wright-Patterson AFB.

The purpose of this research is to identify conditions under which individual operators and teams may be vulnerable to change blindness, the striking phenomenon in which observers often fail to detect changes to objects in visual scenes from one view to the next. This study also aims to determine whether teams are less susceptible to change blindness than individual operators, and to identify strategies used by teams to detect changes.

The time requirement for each volunteer participant is anticipated to be a total of 1 visit of approximately 2 hours. Approximately 150 participants will be enrolled in this study. You will be required to have normal or corrected-to-normal visual acuity to be eligible for participation.

2. **Experimental procedures:** If you decide to participate you will be asked to view tactical displays on a desktop PC with the task of detecting heading, color, and position changes either individually or as part of a team. Testing will last approximately 2 hours and will be done in normal lighting conditions. Normally, you will be seated during the data collection trials. You will be offered a rest period midway through the testing period but can request a break at any time during the experiment. You may withdraw this consent at any time and discontinue further participation in this study without prejudice to your entitlements. Also understand that the medical monitor of this study may terminate your participation in this study if she or he feels this to be in your best interest.

While performing the simulation behavioral and physiological measures will be collected. Specifically, either eye movements, or heart rate, or both will be recorded. An ASL Model 501 head-mounted eye tracker will be used to record your eye movements and an electrocardiogram (ECG) will be used to record your heart rate. None of these methods are physically invasive.

3. **Discomfort and risks:** There is minimal risk and/or discomfort associated with performing this task. Mild postural fatigue and eye strain has been shown to be prevalent in continued computer usage but is normally alleviated with rest breaks. There is a small risk that the leads used for the heart tracing (ECG) may irritate your skin.

4. **Precautions for female participants:** There are no special precautions for female participants.

5. **Benefits:** You are not expected to benefit directly from participation in this research study.

6. **Compensation**: You will receive the standard compensation of $15.00 per hour (cash) as a member of the General Dynamics paid volunteer subject pool.

7. **Alternatives**: Choosing not to participate is an alternative to volunteering for this study.

8. **Entitlements and confidentiality:**

a. Records of your participation in this study may only be disclosed according to federal law, including the Federal Privacy Act, 5 U.S.C. 552a, and its implementing regulations.

b. You understand your entitlements to medical and dental care and/or compensation in the event of injury are governed by federal laws and regulations, and that if you desire further information you may contact the base legal office (88 ABW/JA, 257-6142 for Wright-Patterson AFB). You may contact the medical monitor, Dr. Jeff Bidinger, Maj., USAF, of this research study at (937) 656-5449.

c. If an unanticipated event (medical misadventure) occurs during your participation in this study, you will be informed. If you are not competent at the time to understand the nature of the event, such information will be brought to the attention of your next of kin.

Next of Kin if needed, Name______________________,

Phone#_________________.

d. The decision to participate in this research is completely voluntary on your part. No one has coerced or intimidated you into participating in this program. You are participating because you want to. Todd Nelson, or an associate, has adequately answered any and all questions you have about this study, your participation, and the procedures involved. Todd Nelson can be reached at (937) 255-8803. You

understand that Alison Tollner, or an associate will be available to answer any questions concerning procedures throughout this study. You understand that if significant new findings develop during the course of this research, which may relate to your decision to continue participation, you will be informed. You further understand that you may withdraw this consent at any time and discontinue further participation in this study without prejudice to your entitlements. You also understand that the medical monitor of this study may terminate your participation in this study if she or he feels this to be in your best interest. If you have any questions or concerns about your participation in this study or your rights as a research participant, please contact Major Jeff Bidinger at (937) 656-5449 or jeffrey.bidinger@wpafb.af.mil.

e. You understand that your participation in this study may be photographed, filmed or audio/videotaped. The audio/video data will be used for data analysis, data retrieval and backup purposes only. All audio/video media will be stored in a secure cabinet for up to 5 years in the BMC2 Lab, Bldg 33, WPAFB, OH. You consent to the use of these media for training and data collection purposes and understand that any release of records of your participation in this study may only be disclosed according to federal law, including the Federal Privacy Act, 55 U.S.C. 552a, and its implementing regulations. This means personal information will not be released to unauthorized source without your permission.

f. YOU FULLY UNDERSTAND THAT YOU ARE MAKING A DECISION WHETHER OR NOT TO PARTICIPATE. YOUR SIGNATURE INDICATES THAT YOU HAVE DECIDED TO PARTICIPATE HAVING READ THE INFORMATION PROVIDED ABOVE.

Volunteer

Signature____Date_______________

Volunteer Social Security No. (Optional)________________________________

Advising Investigator Signature ______________________ Date _________________

Witness Signature __________________________________Date _________________

Privacy Act Statement

<u>Authority</u>: We are requesting disclosure of personal information, to include your Social Security Number. Researchers are authorized to collect personal information (including social security numbers) on research participants under The Privacy Act-5 USC 552a, 10 USC 55, 10 USC 8013, 32 CFR 219, 45 CFR Part 46, and EO 9397, November 1943.

<u>Purpose</u>: It is possible that latent risks or injuries inherent in this experiment will not be discovered until some time in the future. The purpose of collecting this information is to aid researchers in locating you at a future date if further disclosures are appropriate.

<u>Routine Uses</u>: Information (including name and SSN) may be furnished to Federal, State

and local agencies for any uses published by the Air Force in the Federal Register, 52 FR 16431, to include, furtherance of the research involved with this study and to provide medical care.

Disclosure: Disclosure of the requested information is voluntary. No adverse action whatsoever will be taken against you, and no privilege will be denied you based on the fact you do not disclose this information. However, your participation in this study may be impacted by a refusal to provide this information.

Individual and Team Susceptibility to Change Blindness – Eye Tracking Study

Participant Biographical Form

Participant Name: ________________________ **Date: ________________**

Age: _____________

Gender: Male Female

Handedness: Left Right Both

Vision: Normal 20/20 Corrected to Normal Deficient

Education level: ________________________________

Total years of military experience: ______________ Rank: ______________

How many hours per week do you play video/computer games?

1.) 0 hrs/week   2.) 1-10 hrs/week   3.) 11-20 hrs/week

4.) 21-30 hrs/week   5.) 30+ hrs/week

Using the following scale, indicate how much you enjoy playing video/computer games?

1	2	3	4	5	6	7
Not at All Enjoy						Very much Enjoy

Appendix B: Task Instructions

Appendix B1: Instructions for Individual Observers

Instructions for change detection task

You will see an image flashing on the screen with either 6 or 48 colored squares. On some trials, one of the squares will change in position. Your job is to indicate (1) if one of the squares changes in position during a trial and (2) which square changed in position. Please press “YES” to indicate that you saw one of the squares change in position and “NO” if you think that none of the squares changed in position. On trials in which you respond “YES,” that a change in position occurred, you will be asked to indicate which square changed in position by using the mouse to select it. Please respond as quickly AND accurately as you can. Press the space bar to move on to the next trial.

The task consists of 2 sets of trials, one set with only 6 squares on the screen and one set with 48 squares on the screen. The number of squares on the screen will not vary within a set of trials. You will start with either one or the other set and you will be given a break between sets of trials. AT THE END OF EACH SET OF TRIALS, YOU WILL BE ASKED TO COMPLETE THE NASA TASK LOAD INDEX QUESTIONNAIRE in which you will be asked to rate the demands of the task, your performance, and the levels of effort you exerted and frustration you experienced during the task. You will also be asked to complete the SART questionnaire in which you will be required to rate other dimensions of the task. Those will be described on the questionnaire. Please let the experimenter know when a message appears on the screen indicating that it is time to fill out these questionnaires.

Please use your LEFT hand for the YES and NO keys and your RIGHT hand to select with the mouse.

Appendix B2: Instructions for Dyads

Instructions for change detection task

You and your teammate will see an image flashing on the screen with either 6 or 48 colored squares. On some trials, one of the squares will change in position. Your job is to indicate (1) if one of the squares changes in position during a trial and (2) which square changed in position. Please press “YES” to indicate that you saw one of the squares change in position and “NO” if you think that none of the squares changed in position. On trials in which you respond “YES,” that a change in position occurred, you will be asked to indicate which square changed in position by using the mouse to select it. Please respond as quickly AND accurately as you can. To continue to the next trial, the person on the left will have to press the space bar.

During the first part of the experiment, you and your teammate will not be allowed to communicate in anyway during the task or between blocks of trials. The first response given during a trial – whether from you or your teammate – will be the recorded response for that trial. During the second part of the experiment, you and your teammate will be allowed to communicate during the experiment and are encouraged to come up with a strategy that will help you detect changes in the images on the screen. The experimenter will indicate the beginning of the second part of the experiment.

The task consists of 2 sets of trials, one set with only 6 squares on the screen and one set with 48 squares on the screen. The number of squares on the screen will not vary within a set of trials. You will start with either one or the other set and you will be given a break between sets of trials. AT THE END OF EACH SET OF TRIALS, YOU WILL BE ASKED

TO COMPLETE THE NASA TASK LOAD INDEX QUESTIONNAIRE in which you will be asked to rate the demands of the task, your performance, and the levels of effort you exerted and frustration you experienced during the task. You will also be asked to complete the SITUATION AWARENESS QUESTIONNAIRE and the TEAM WORKLOAD SCALE in which you will be required to rate other dimensions of the task. Those will be described on the questionnaires. Please let the experimenter know when a message appears on the screen indicating that it is time to fill out these questionnaires.

Please use your LEFT hand for the YES and NO keys and your RIGHT hand to select with the mouse.

Appendix C: Questionnaires

Appendix C1: NASA TLX

Change Blindness Study

Participant I.D.: ___________Date: _____________ Trial #: _____________

Condition:._______________

NASA TLX (Task Load Index) Rating Scales

Instructions: For each of the scales presented below, ***please write the number*** *that matches your experience with what you just completed. Please note that the "Performance" scale goes from "good" on the left to "poor" on the right.*

Mental Demand

Low 0 25 50 75 100 **High**

Performance

Good 0 25 50 75 100 **Poor**

Physical Demand

Low 0 25 50 75 100 **High**

Effort

Low 0 25 50 75 100 **High**

Temporal Demand

Low 0 25 50 75 100 **High**

Frustration

Low 0 25 50 75 100 **High**

Appendix C2: Team Work Load Scale

Communication Demand—

How much communication was required between you and other team members directly or via the computer?

---
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Low High

Did requesting and transferring information consume a little of your time or a lot?

---
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 A
Little A Lot

Was it easy or demanding?

---
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Easy Demanding

Monitoring Demand—

How much monitoring of people did the task require?

---
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Low High

Was attending to others (directly or through the computer) easy or demanding?

---
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Easy Demanding

Was attending to others (directly or through the computer) infrequent or continuous?

---
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Infrequent Continuous

Control Demand—

How much correction of others did the task require?

---
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 A
Little A Lot

Was correcting other people simple or complex?

---
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Simple Complex

Was correcting other people infrequent or continuous?

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Infrequent Continuous

Coordination Demand—

How much correction or adjustment of your own actions was required in order to coordinate with others?

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
A Little A Lot

Was adjusting your actions to improve coordination simple or complex?

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Simple Complex

Was adjusting your actions to improve coordination infrequent or continuous?

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Infrequent Continuous

Leadership Demand—

How much leadership was required of you and/or your team members?

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
A Little A Lot

Was being a leader easy or demanding?

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Easy Demanding

Was a leader needed infrequently or frequently?

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Infrequently Frequently

Appendix C3: Exit Questionnaire

Change Blindness Study

Participant I.D.: __________ **Date: _____________** **Trial #: _____________**

Condition:.___________

Exit Questionnaire

Did you adopt any strategies to aid yourself in detecting changes? If so, what were they?

_____

____

____

______

At any point throughout the experiment did you have a feeling that something was changing but could not identify it?

____

____

___

When a change was noticed did you respond immediately or did you wait for another iteration to occur so you could see the change again before responding?

Did you notice any patterns in the changes?

How confident were you when it came to detecting changes? Please circle your response by using a 1-10 rating scale (1: not confident, 10: very confident).

1	2	3	4	5	6	7	8	9	10
not confident									very confident

Any additional comments?

Appendix D: Analysis of Accuracy from *Individuals* and *Comm Dyads* With and Without Participant Q

INCLUDING PARTICIPANT Q's DATA

Effects coding used for categorical variables in model.
Categorical values encountered during processing are:
indiv_v_comm $ (2 levels)
individuals, comm dyads
7 case(s) deleted due to missing data.
Number of cases processed: 16
Dependent variable means

AC_6LG	AC_6SM	AC_6NC	AC_48LG	AC_48SM	AC_48NC
0.995	0.896	0.995	0.896	0.708	0.999

Repeated measures factors and levels
Dependent Variables

Within factor	1	2	3	4	5	6
set size	1.000	1.000	1.000	2.000	2.000	2.000
change magn	1.000	2.000	3.000	1.000	2.000	3.000

Univariate and multivariate Repeated Measures Analysis
Between Subjects

Source	SS	df	MS	F	P
indiv_v_ comm $	0.196	1	0.196	5.622	0.033
Error	0.488	14	0.035		

Within Subjects

Source	SS	df	MS	F	P	G-G	H-F
setsize	0.177	1	0.177	9.780	0.007	.	.
setsize* indiv_v_comm $	0.081	1	0.081	4.448	0.053	.	.
Error	0.254	14	0.018				

Greenhouse-Geisser Epsilon: .
Huynh-Feldt Epsilon : .

Source	SS	df	MS	F	P	G-G	H-F
changemagn	0.552	2	0.276	17.431	0.000	0.000	0.000
changemagn* indiv_v_comm $	0.218	2	0.109	6.875	0.004	0.008	0.005
Error	0.444	28	0.016				

Greenhouse-Geisser Epsilon: 0.7713
Huynh-Feldt Epsilon : 0.9103

Source	SS	df	MS	F	P	G-G	H-F
setsize*changemagn	0.122	2	0.061	5.538	0.009	0.012	0.009
setsize*changemagn* indiv_v_comm $	0.066	2	0.033	2.999	0.066	0.071	0.066
Error	0.309	28	0.011				

Greenhouse-Geisser Epsilon: 0.9197
Huynh-Feldt Epsilon : 1.0000
Multivariate repeated measures analysis
Test of: changemagn

Statistic	Value	Hypoth. df	Error df	F	P
Wilks' Lambda	0.381	2	13	10.548	0.002
Pillai Trace	0.619	2	13	10.548	0.002
H-L Trace	1.623	2	13	10.548	0.002

Test of: changemagn* indiv_v_comm $

Statistic	Value	Hypoth. df	Error df	F	P
Wilks' Lambda	0.600	2	13	4.340	0.036
Pillai Trace	0.400	2	13	4.340	0.036
H-L Trace	0.668	2	13	4.340	0.036

Test of: setsize*changemagn

Statistic	Value	Hypoth. df	Error df	F	P
Wilks' Lambda	0.509	2	13	6.258	0.012
Pillai Trace	0.491	2	13	6.258	0.012
H-L Trace	0.963	2	13	6.258	0.012

Test of: setsize*changemagn* indiv_v_comm $

Statistic	Value	Hypoth. df	Error df	F	P
Wilks' Lambda	0.667	2	13	3.248	0.072
Pillai Trace	0.333	2	13	3.248	0.072
H-L Trace	0.500	2	13	3.248	0.072

EXCLUDING PARTICIPANT Q's DATA

Effects coding used for categorical variables in model.
Categorical values encountered during processing are:
indiv_v_comm $ (2 levels)
individuals, comm dyads
7 case(s) deleted due to missing data.
Number of cases processed: 15

Dependent variable means

AC_6LG	AC_6SM	AC_6NC	AC_48LG	AC_48SM	AC_48NC
0.994	0.917	0.994	0.939	0.750	0.999

Repeated measures factors and levels
Dependent Variables

Within factor	1	2	3	4	5	6
set size	1.000	1.000	1.000	2.000	2.000	2.000
change magn	1.000	2.000	3.000	1.000	2.000	3.000

Univariate and multivariate Repeated Measures Analysis
Between Subjects

Source	SS	df	MS	F	P
indiv_v_comm $	0.090	1	0.090	7.551	0.017
Error	0.154	13	0.012		

Within Subjects

Source	SS	df	MS	F	P	G-G	H-F
setsize	0.109	1	0.109	10.906	0.006	.	.
setsize* indiv_v_comm $	0.039	1	0.039	3.863	0.071	.	.
Error	0.130	13	0.010				

Greenhouse-Geisser Epsilon: .
Huynh-Feldt Epsilon : .

Source	SS	df	MS	F	P	G-G	H-F
changemagn	0.418	2	0.209	20.581	0.000	0.000	0.000
changemagn* indiv_v_comm $	0.149	2	0.074	7.321	0.003	0.014	0.011
Error	0.264	26	0.010				

Greenhouse-Geisser Epsilon: 0.5732
Huynh-Feldt Epsilon : 0.6409

Source	SS	df	MS	F	P	G-G	H-F
setsize*changemagn	0.103	2	0.051	6.476	0.005	0.017	0.013
setsize*changemagn* indiv_v_comm $	0.057	2	0.029	3.614	0.041	0.068	0.061
Error	0.206	26	0.008				

Greenhouse-Geisser Epsilon: 0.6234
Huynh-Feldt Epsilon : 0.7105
Multivariate repeated measures analysis
Test of: changemagn

Statistic	Value	Hypoth. df	Error df	F	P
Wilks' Lambda	0.247	2	12	18.258	0.000
Pillai Trace	0.753	2	12	18.258	0.000
H-L Trace	3.043	2	12	18.258	0.000

Test of: changemagn* indiv_v_comm $

Statistic	Value	Hypoth. df	Error df	F	P
Wilks' Lambda	0.603	2	12	3.946	0.048
Pillai Trace	0.397	2	12	3.946	0.048
H-L Trace	0.658	2	12	3.946	0.048

Test of: setsize*changemagn

Statistic	Value	Hypoth. df	Error df	F	P
Wilks' Lambda	0.437	2	12	7.725	0.007
Pillai Trace	0.563	2	12	7.725	0.007
H-L Trace	1.287	2	12	7.725	0.007

Test of: setsize*changemagn* indiv_v_comm $

Statistic	Value	Hypoth. df	Error df	F	P
Wilks' Lambda	0.675	2	12	2.884	0.095
Pillai Trace	0.325	2	12	2.884	0.095
H-L Trace	0.481	2	12	2.884	0.095

22 cases and 50 variables processed and saved.

Appendix E: Data

Appendix E1: Performance and fixations data, dyads

subject$	RT_6LGn	RT_6SMn	RT_6NCn	RT_48LGn	RT_48SMn	RT_48NCn	RT_6LGc	RT_6SMc	RT_6NCc	RT_48LGc	RT_48SMc	RT_48NCc
dyad_3	1336	2539	4554.5	7554	4828	15319.5	1336.5	1773.5	3062.5	4351.5	4453	10812
dyad_4	859.5	1360	3273.5	3484.5	5406.5	16922	1282	1430	4648.5	2867.5	4328	13461.5
dyad_5	1344	1437.5	1656	3390	3922	5406	906.5	2484	3109.5	3461	3656	6601
dyad_6	1336	1500	2508	4274	2453	10569.5	1157	1422	2898.5	4804.5	5563	11539
dyad_8	1343	1702	3849.5	4616	2457	8402	1070.5	1336	4921.5	2430	2671.5	10125
dyad_9	813	1188	2468.5	4633	4532	11430	1320.5	703	4976.5	2407	2484	9969.5
dyad_11	1384	1564	2800	4839.5	4190	9335.5	1141.5	1602.5	1517	4809	2866.5	7597.5

subject$	AC_6LGn	AC_6SMn	AC_6NCn	AC_48LGn	AC_48SMn	AC_48NCn	AC_6LGc	AC_6SMc	AC_6NCc	AC_48LGc	AC_48SMc	AC_48NCc
dyad_3	1	1	1	0.83333333	1	1	1	1	1	1	0.91666667	1
dyad_4	1	1	1	1	0.83333333	1	0.91666667	1	1	1	1	1
dyad_5	0.91666667	0.66666667	1	0.58333333	0.41666667	0.91666667	1	0.916667	1	0.83333333	0.75	0.999
dyad_6	1	1	1	1	0.91666667	1	1	0.916667	1	1	0.91666667	1
dyad_8	1	1	1	0.91666667	0.91666667	1	1	1	1	1	1	1
dyad_9	1	1	1	1	0.75	1	1	1	1	0.91666667	0.91666667	1
dyad_11	1	0.83333333	0.91666667	0.83333333	0.58333333	1	1	0.833333	0.999	0.91666667	0.83333333	1

subject$	fx_lgch6n	fx_smch6n	fx_noch6n	fx_lgch48n	fx_smch48n	fx_noch48n	fx_lgch6c	fx_smch6c	fx_noch6c	fx_lgch48c	fx_smch48c	fx_noch48c
dyad_3	10.25	15.5	22.833333	31.916667	25.1666667	77.25	10	13.583333	18.5	26.25	24.75	55.25
dyad_4	7.833333	9.4166667	15.583333	23.083333	27	87	9.166667	12.083333	26.5	18.416667	39.1666667	70.083333
dyad_5	7	5.0833333	9.4166667	12.5	9.75	27.75	7.666667	10.5	15.75	19.583333	15.1666667	33.416667
dyad_6	7.583333	9.5833333	13.583333	22.833333	16.8333333	46.833333	9.416667	9.0833333	16.666667	24.166667	27.25	51.916667
dyad_8	10.75	12.333333	25.166667	24.083333	18	97.25	9.916667	9.4166667	22.583333	20.416667	20.3333333	84.166667
dyad_9	6.583333	8.4166667	10.5	19.166667	17.8333333	43.333333	8.333333	8.75	17.583333	14	15.1666667	45.75
dyad_11	8.833333	8.0833333	10.75	19.916667	15.9166667	48.833333	9.5	9.1666667	17.25	21.75	18	41.25

Appendix E2: Performance and fixations data, Individuals vs Dyads

subject$	ivnocom$	ivcom$	fx_lgch6	fx_smch6	fx_noch6	fx_lgch48	fx_smch48	fx_noch48
indiv_2	indiv	indiv	5.833333333	4.25	8.916666667	20	8.5	33.25
indiv_3	indiv	indiv	8.166666667	9.416666667	17.33333333	17.33333333	18.83333333	31.33333333
indiv_7	indiv	indiv	9.083333333	11.5	21.58333333	27	7.5	59
indiv_10	indiv	indiv	7.5	8.666666667	17.41666667	12.58333333	14.58333333	32.16666667
indiv_14	indiv	indiv	5.25	3.5	5.833333333	14.41666667	7.333333333	27.16666667
indiv_15	indiv	indiv	6	2.833333333	7.083333333	17.33333333	9.833333333	34.58333333
indiv_16	indiv	indiv	3.916666667	4.75	10.58333333	10.91666667	14.83333333	37.08333333
indiv_20	indiv	indiv	12.75	20.16666667	26.58333333	44.75	19.66666667	86.41666667
indiv_21	indiv	indiv	3.75	2.166666667	5.083333333	1.666666667	1.333333333	13.66666667
dyad_3	no comm		10.25	15.5	22.83333333	31.91666667	25.16666667	77.25
dyad_4	no comm		7.833333333	9.416666667	15.58333333	23.08333333	27	87
dyad_5	no comm		7	5.083333333	9.416666667	12.5	9.75	27.75
dyad_6	no comm		7.583333333	9.583333333	13.58333333	22.83333333	16.83333333	46.83333333
dyad_8	no comm		10.75	12.33333333	25.16666667	24.08333333	18	97.25
dyad_9	no comm		6.583333333	8.416666667	10.5	19.16666667	17.83333333	43.33333333
dyad_11	no comm		8.833333333	8.083333333	10.75	19.91666667	15.91666667	48.83333333
dyad_3		comm	10	13.58333333	18.5	26.25	24.75	55.25
dyad_4		comm	9.166666667	12.08333333	26.5	18.41666667	39.16666667	70.08333333
dyad_5		comm	7.666666667	10.5	15.75	19.58333333	15.16666667	33.41666667
dyad_6		comm	9.416666667	9.083333333	16.66666667	24.16666667	27.25	51.91666667
dyad_8		comm	9.916666667	9.416666667	22.58333333	20.41666667	20.33333333	84.16666667
dyad_9		comm	8.333333333	8.75	17.58333333	14	15.16666667	45.75
dyad_11		comm	9.5	9.166666667	17.25	21.75	18	41.25

Appendix E3: Re-scanning and cross-scanning data, Dyads

cr_noch48	cr_smch48	cr_lgch48	cr_noch6n	cr_smch6n	cr_lgch6n	cr_noch48n	cr_smch48n	cr_lgch48n
0.667	0.091	0.083	1.000	0.917	0.583	1.000	0.333	0.600
1.000	0.417	0.250	0.917	0.417	0.333	1.000	0.700	0.500
0.167	0.000	0.100	0.500	0.375	0.455	1.000	0.800	0.714
1.000	0.455	0.250	1.000	0.500	0.250	1.000	0.364	0.583
1.000	0.083	0.000	0.917	0.500	0.667	1.000	0.364	0.545
0.667	0.000	0.000	0.667	0.500	0.250	1.000	0.333	0.417
0.500	0.100	0.091	0.727	0.600	0.583	1.000	0.857	0.900

Appendix E4: Re-scanning and cross-scanning data, Individuals vs Dyads

Group$	ivnocom$	ivcom$	re_noch6	re_smch6	re_lgch6	re_noch48	re_smch48	re_lgch48
indiv_2	indiv	indiv	0.917	0.727	0.833	1.000	1.000	1.000
indiv_3	indiv	indiv	1.000	1.000	0.833	1.000	1.000	0.917
indiv_7	indiv	indiv	1.000	1.000	1.000	1.000	1.000	1.000
indiv_10	indiv	indiv	1.000	1.000	1.000	1.000	1.000	0.833
indiv_14	indiv	indiv	0.667	0.600	0.917	1.000	1.000	1.000
indiv_15	indiv	indiv	0.909	0.833	0.917	1.000	1.000	1.000
indiv_16	indiv	indiv	0.917	0.750	0.583	1.000	0.818	1.000
indiv_20	indiv	indiv	1.000	1.000	0.917	1.000	1.000	1.000
indiv_21	indiv	indiv	0.667	0.714	0.583	1.000	1.000	1.000
dyad_3_1	no Comm		1.000	1.000	0.833	1.000	1.000	0.900
dyad_4_1	no Comm		1.000	0.667	0.583	1.000	1.000	1.000
dyad_5_1	no Comm		0.333	0.250	0.364	1.000	1.000	0.714
dyad_6_1	no Comm		1.000	0.750	0.583	1.000	0.909	0.917
dyad_8_1	no Comm		1.000	0.917	0.750	1.000	1.000	1.000
dyad_9_1	no Comm		0.833	0.833	0.667	1.000	0.889	0.917
dyad_11_1	no Comm		0.818	0.800	0.917	1.000	1.000	1.000
dyad_3_2	no Comm		1.000	0.667	0.750	1.000	0.917	1.000
dyad_4_2	no Comm		0.833	0.750	0.667	1.000	1.000	0.917
dyad_5_2	no Comm		0.917	0.875	0.727	1.000	1.000	0.857
dyad_6_2	no Comm		0.917	0.750	0.583	1.000	0.909	0.833
dyad_8_2	no Comm		0.917	0.917	0.583	1.000	1.000	0.909
dyad_9_2	no Comm		0.833	0.583	0.333	1.000	0.778	0.750
dyad_11_2	no Comm		0.818	0.600	0.500	1.000	1.000	1.000
dyad_3_1		Comm	1.000	1.000	1.000	1.000	1.000	1.000
dyad_4_1		Comm	1.000	0.917	0.727	1.000	1.000	1.000
dyad_5_1		Comm	0.833	0.727	0.333	0.917	0.778	1.000
dyad_6_1		Comm	1.000	0.818	0.833	1.000	1.000	0.917
dyad_8_1		Comm	1.000	0.917	1.000	1.000	0.833	1.000
dyad_9_1		Comm	1.000	0.667	0.667	1.000	0.818	0.818
dyad_11_1		Comm	1.000	0.900	0.750	1.000	1.000	0.909
dyad_3_2		Comm	0.917	0.917	0.750	1.000	1.000	0.917
dyad_4_2		Comm	1.000	0.917	0.818	1.000	0.917	0.750
dyad_5_2		Comm	1.000	0.909	0.667	1.000	1.000	1.000
dyad_6_2		Comm	1.000	0.364	0.750	1.000	1.000	0.917
dyad_8_2		Comm	0.917	0.583	0.500	1.000	0.833	0.833
dyad_9_2		Comm	0.833	0.583	0.500	1.000	0.909	0.636
dyad_11_2		Comm	0.833	0.600	0.750	1.000	0.900	1.000

Appendix E5: workload data, Dyads

SUBJECT$	RTLX 6no	RTLX 48no	RTLX 6c	RTLX 48c	NNI_6n	NNI_48n	NNI_6c	NNI_48c	HR_6n	HR_48n	HR_6c	HR_48c
dyad_3	45.417	56.250	38.750	48.333	0.912	0.877	0.911	0.769	787.082	773.035	801.725	797.021
dyad_4	37.917	44.583	29.583	46.667	0.834	0.872	0.833	0.734	1051.045	972.390	939.518	871.052
dyad_5	56.667	55.000	35.833	44.583	0.854	0.877	0.842	0.856	773.473	782.442	797.412	785.669
dyad_6	22.083	36.250	23.333	26.667	0.859	0.877	0.842	0.732	846.751	813.826	678.296	699.684
dyad_8	28.750	35.000	24.167	25.833	0.839	0.817	0.799	0.786	919.643	859.103	810.962	823.439
dyad_9	35.667	35.250	28.750	29.333	0.875	0.910	0.859	0.876	907.441	913.989	949.492	949.265
dyad_11	52.500	56.667	53.750	59.583	0.838	0.825	0.860	0.826	746.317	756.185	885.153	864.656

Appendix E6: workload data, Individuals vs Dyads

SUBJECT$	ivnocom$	ivcom$	RTLX_6	RTLX_48	NNI_6	NNI_48	HR_6	HR_48
indiv_2	indiv	indiv	35.000	60.833	0.849	0.835		
indiv_3	indiv	indiv	35.000	46.667	0.798	0.882		
indiv_7	indiv	indiv	20.833	60.833	0.841	0.879		
indiv_10	indiv	indiv	24.167	51.667	0.874	0.896	711.325	730.321
indiv_14	indiv	indiv	37.500	55.000	0.916	0.818	805.007	805.993
indiv_15	indiv	indiv	33.333	57.500	0.761	0.879	1094.387	1018.252
indiv_16	indiv	indiv	23.167	18.667	0.954	0.916	995.104	891.658
indiv_20	indiv	indiv	32.500	32.500	0.813	0.813	931.016	924.963
indiv_21	indiv	indiv	33.333	25.833	0.899	0.913	749.427	796.622
dyad_3	nocomm		45.417	56.250	0.912	0.877	787.082	773.035
dyad_4	nocomm		37.917	44.583	0.834	0.872	1051.045	972.390
dyad_5	nocomm		56.667	55.000	0.854	0.877	773.473	782.442
dyad_6	nocomm		22.083	36.250	0.859	0.877	846.751	813.826
dyad_8	nocomm		28.750	35.000	0.839	0.817	919.643	859.103
dyad_9	nocomm		35.667	35.250	0.875	0.910	907.441	913.989
dyad_11	nocomm		52.500	56.667	0.838	0.825	746.317	756.185
dyad_3		comm	38.750	48.333	0.911	0.769	801.725	797.021
dyad_4		comm	29.583	46.667	0.833	0.734	939.518	871.052
dyad_5		comm	35.833	44.583	0.842	0.856	797.412	785.669
dyad_6		comm	23.333	26.667	0.842	0.732	678.296	699.684
dyad_8		comm	24.167	25.833	0.799	0.786	810.962	823.439
dyad_9		comm	28.750	29.333	0.859	0.876	949.492	949.265
dyad_11		comm	53.750	59.583	0.860	0.826	885.153	864.656

Appendix E7: NNI correlations to heart rate and RTLX

SUBJECT$	SETSIZE$	SUBJECT	SETSIZE	HR	RTLX	NNI
indiv_2	6.000	1	1		35	0.8490266
indiv_3	6.000	1	1		35	0.7980111
indiv_7	6.000	1	1		20.833333	0.8412377
indiv_10	6.000	1	1		24.166667	0.8741777
indiv_14	6.000	1	1	711.325	37.5	0.916476
indiv_15	6.000	1	1	805.00725	33.333333	0.7613795
indiv_16	6.000	1	1	1094.3873	23.166667	0.9535818
indiv_20	6.000	1	1	995.10427	32.5	0.8133558
indiv_21	6.000	1	1	931.01575	33.333333	0.8990996
dyad3_1n	6.000	2	1	791.17452	46.666667	0.840472
dyad4_1n	6.000	2	1	858.6902	23.333333	0.7529475
dyad5_1n	6.000	2	1	909.31169	60	0.8183159
dyad6_1n	6.000	2	1	819.10781	24.166667	0.8208879
dyad8_1n	6.000	2	1	906.47603	35.833333	0.8659857
dyad9_1n	6.000	2	1	1186.1077	31.333333	0.8880612
dyad11_1n	6.000	2	1	979.86194	47.5	0.7996474
dyad3_2n	6.000	2	1	707.67901	44.166667	0.8272924
dyad4_2n	6.000	2	1	715.47319	52.5	0.9552916
dyad5_2n	6.000	2	1	1192.7787	53.333333	0.9003962
dyad6_2n	6.000	2	1	727.83828	20	0.8574745
dyad8_2n	6.000	2	1	787.02663	21.666667	0.8836778
dyad9_2n	6.000	2	1	653.17839	40	0.7882323
dyad11_2n	6.000	2	1	835.01987	57.5	1.024145
dyad3_1c	6.000	3	1	796.20755	46.666667	0.834237
dyad4_1c	6.000	3	1	856.22963	21.666667	0.8014305
dyad5_1c	6.000	3	1	823.72911	40	0.8503429
dyad6_1c	6.000	3	1	841.94574	22.5	0.7449266
dyad8_1c	6.000	3	1	846.79433	25.833333	0.8908347
dyad9_1c	6.000	3	1	1306.1859	28.333333	0.8828942
dyad11_1c	6.000	3	1	941.66791	36.666667	0.864321
dyad3_2c	6.000	3	1	696.42582	25	0.8320357
dyad4_2c	6.000	3	1	747.22006	21.666667	0.8832723
dyad5_2c	6.000	3	1	1055.3074	33.333333	0.8336229
dyad6_2c	6.000	3	1	752.87847	20	0.8522809
dyad8_2c	6.000	3	1	775.12945	20	0.8268384
dyad9_2c	6.000	3	1	592.79744	28.333333	0.8376425
dyad11_2c	6.000	3	1	828.63722	63.333333	0.957224
indiv_2	48.000	1	2		60.833333	0.834775
indiv_3	48.000	1	2		46.666667	0.8823834
indiv_7	48.000	1	2		60.833333	0.8791435
indiv_10	48.000	1	2		51.666667	0.8961612
indiv_14	48.000	1	2	730.32056	55	0.8175046
indiv_15	48.000	1	2	805.99267	57.5	0.8792771
indiv_16	48.000	1	2	1018.2524	18.666667	0.9159006
indiv_20	48.000	1	2	891.65825	32.5	0.8133558
indiv_21	48.000	1	2	924.96273	25.833333	0.9133735
dyad3_1n	48.000	2	2	877.23383	49.166667	0.8947366
dyad4_1n	48.000	2	2	830.08537	43.333333	0.8008798
dyad5_1n	48.000	2	2	898.69048	46.666667	0.8251116
dyad6_1n	48.000	2	2	832.87302	38.333333	0.8924323
dyad8_1n	48.000	2	2	851.64796	45	0.8250864
dyad9_1n	48.000	2	2	1104.326	28.833333	0.8250864
dyad11_1n	48.000	2	2	996.98675	55	0.8857912
dyad3_2n	48.000	2	2	716.00947	63.333333	0.849566
dyad4_2n	48.000	2	2	715.98494	45.833333	0.9528295
dyad5_2n	48.000	2	2	1046.0886	63.333333	0.9282681
dyad6_2n	48.000	2	2	732.01198	34.166667	0.7422238
dyad8_2n	48.000	2	2	776.0031	25	0.9943563
dyad9_2n	48.000	2	2	613.87942	41.666667	0.8251701
dyad11_2n	48.000	2	2	830.99119	58.333333	0.8691516
dyad3_1c	48.000	3	2	823.75147	48.333333	0.8391671
dyad4_1c	48.000	3	2	847.16507	46.666667	0.8577682
dyad5_1c	48.000	3	2	825.975	35.833333	0.7840848
dyad6_1c	48.000	3	2	834.37132	30.833333	0.7932671
dyad8_1c	48.000	3	2	870.04936	31.666667	0.8410208
dyad9_1c	48.000	3	2	1276.1433	20.333333	0.8393137
dyad11_1c	48.000	3	2	903.29426	54.166667	0.7783499
dyad3_2c	48.000	3	2	688.61866	48.333333	0.6282463
dyad4_2c	48.000	3	2	746.87648	46.666667	0.8536838
dyad5_2c	48.000	3	2	916.12911	53.333333	0.6798736
dyad6_2c	48.000	3	2	736.96707	22.5	0.7791402
dyad8_2c	48.000	3	2	776.82871	20	0.9109248
dyad9_2c	48.000	3	2	622.38728	38.333333	0.8136616
dyad11_2c	48.000	3	2	826.01828	65	0.7599139

Appendix E8: Regression model of change detection with multiple observers

SUBJECT	IVNOCOM	SETSIZE$	HR	NNI	RTLX	CONF	RESCAN	CRSCAN	FIX	RT	ACC
cbit3	nocomma	six	749.4268	0.833136	45.41667	8.5	0.875	0.833333	16.19444	2809.833	1
cbit4	nocomma	six	787.0817	0.842351	37.91667	9	0.75	0.555556	10.94444	1831	1
cbit5	nocomma	six	1051.045	0.841983	56.66667	8	0.577652	0.443182	7.166667	1479.167	0.861111
cbit6	nocomma	six	773.473	0.798604	22.08333	9.5	0.763889	0.583333	10.25	1781.333	1
cbit8	nocomma	six	846.7513	0.858837	28.75	8	0.847222	0.694444	16.08333	2298.167	1
cbit9	nocomma	six	919.643	0.860268	35.66667	8.5	0.680556	0.472222	8.5	1489.833	1
cbit3	comm ave	six	746.3167	0.833882	35.83333	9	0.930556	0.333333	14.02778	2057.5	1
cbit4	comm ave	six	801.7248	0.85412	21.66667	9	0.896465	0.623737	15.91667	2453.5	0.972222
cbit5	comm ave	six	939.5183	0.859356	36.66667	9	0.74495	0.257576	11.30556	2166.667	0.972222
cbit6	comm ave	six	797.4121	0.839181	21.25	9.5	0.794192	0.800505	11.72222	1825.833	0.972222
cbit8	comm ave	six	810.9619	0.874832	22.91667	8.5	0.819444	0.416667	13.97222	2442.667	1
cbit9	comm ave	six	949.4917	0.838147	28.33333	9.5	0.708333	0.5	11.55556	2333.333	1
cbit11	comm ave	six	885.1526	0.911896	50	9	0.805556	0.672222	11.97222	1420.333	0.944111
cbit3	nocomma	forty-eight	796.6217	0.733707	56.25	8.5	0.969444	0.644444	44.77778	9233.833	0.944444
cbit4	nocomma	forty-eight	773.0352	0.855726	44.58333	9	0.986111	0.733333	45.69444	8604.333	0.944444
cbit5	nocomma	forty-eight	972.3896	0.731979	55	8	0.928571	0.838095	16.66667	4239.333	0.638889
cbit6	nocomma	forty-eight	782.4425	0.786204	36.25	9.5	0.92803	0.64899	28.83333	5765.5	0.972222
cbit11	nocomma	forty-eight	913.989	0.769132	56.66667	7.5	1	0.919048	28.22222	6121.667	0.805556
cbit3	comm ave	forty-eight	756.1851	0.872151	48.33333	9	0.986111	0.280303	35.41667	6538.833	0.972222
cbit5	comm ave	forty-eight	871.0521	0.87669	44.58333	9	0.949074	0.088889	22.72222	4572.667	0.861111
cbit9	comm ave	forty-eight	949.2653	0.825128	29.33333	9.5	0.863636	0.222222	24.97222	4953.5	0.944444
cbit11	comm ave	forty-eight	864.6563	0.877471	59.58333	9	0.968182	0.230303	27	5091	0.916667